OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: November 29, 2004, 13:45:01 ; Search time 400 Seconds
                 (without alignments)
                 354.336 Million cell updates/sec

Title: US-10-073-064-2
Perfect score: 27
Sequence: 1 GGTGACTCACCGGGTGAACGGGGGGCATT 27
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_23Sep04:*
                 1: geneseqn1980s:*
                 2: geneseqn1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002as:*
                 7: geneseqn2002bs:*
                 8: geneseqn2003as:*
                 9: geneseqn2003bs:*
                 10: geneseqn2003cs:*
                 11: geneseqn2003ds:*
                 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	27	100.0	27	2	AAQ42309	Aaq42309 Sequence
2	27	100.0	27	2	AAX88057	Aax88057 Vitamin D
3	25	92.6	57	4	AAF85650	Aaf85650 Human ost
4	23	85.2	53	2	AAV18523	Aav18523 Primer VD
5	23	85.2	53	2	AAX16604	Aax16604 Human vit
6	23	85.2	57	4	AAF85650	Aaf85650 Human ost
7	23	85.2	57	4	AAF85639	Aaf85639 Human ost
8	23	85.2	57	4	AAF85639	Aaf85639 Human ost
9	23	85.2	301	2	AAQ79120	Aaq79120 Osteocalc
10	23	85.2	565	8	ABZ58131	Abz58131 Human ost
11	23	85.2	597	5	AAT90018	Aat90018 Bases - 57
12	23	85.2	597	5	AAD14384	Aad14384 Human ost
13	23	85.2	2650	8	AAD48137,	Aad48137 Human ple
14	23	85.2	2650	12	ADN06011	Adn06011 Antipsori
15	23	85.2	2657	8	ABZ34742	Abz34742 Coding se
16	23	85.2	2657	8	ACF12925	Acf12925 Human cer
17	23	85.2	2860	12	ADQ24020	Adq24020 Human sof
18	23	85.2	167932	10	ADL13501	Adl13501 Osteoarth
19	20.2	74.8	13013	10	ADJ72219	Adj72219 Streptomy
20	20.2	74.8	13020	6	ABQ78873	Abq78873 S. roseos
21	19.8	73.3	184	8	ABZ58130	Abz58130 Human ost

22	19.6	72.6	13203	4	ABL30124	Abl30124 Drosophil
23	19	70.4	6846	4	ABL18560	Abl18560 Drosophil
24	19	70.4	12631	4	ABL17964	Abl17964 Drosophil
25	18.2	67.4	132	12	ACH86244	Ach86244 Human gen
26	18.2	67.4	545	12	ACH72488	Ach72488 Human gen
27	18.2	67.4	721	14	AAI59837	Aai59837 Human pol
28	18.2	67.4	995	6	AAD26472	Aad26472 Human kin
29	18.2	67.4	1223	6	AAS27222	Aas27222 cDNA enco
30	18.2	67.4	1223	10	ADB93400	Adb93400 Human CDN
31	18.2	67.4	1440	10	ABV75301	Abv75301 Novel hum
32	18.2	67.4	1452	10	AAD14324	Aad14324 Human 278
33	18.2	67.4	1452	10	AAD61170	Aad61170 Human 278
34	18.2	67.4	1579	10	ADA53813	Ada53813 Human cod
35	18.2	67.4	1581	8	AAD53173	Aad53173 Human kin
36	18.2	67.4	1619	10	ADB63503	Adb63503 Human CDN
37	18.2	67.4	1623	10	ADC10109	Adc10109 Human CDN
38	18.2	67.4	1778	4	AAI58051	Aai58051 Human NOV
39	18.2	67.4	1967	10	ADC10107	Adc10107 Human NOV
40	18.2	67.4	5340	6	ABL32234	Abl32234 Human imm
41	18	66.7	18	8	ABZ58132	Abz58132 Human ost
42	18	66.7	624	11	ACH96158	Ach96158 Klebsiell
43	18	66.7	724	4	AAH08666	Aah08666 Human CDN
44	18	66.7	768	11	ACH96135	Ach96135 Klebsiell
45	18	66.7	1063	8	ACA35059	Aca35059 Prokaryot

ALIGNMENTS

RESULT 1
AAQ42309
ID   AAQ42309 standard; cDNA; 27 BP.
XX
AC   AAQ42309;
XX
DT   25-MAR-2003 (revised)
DT   09-JAN-2003 (revised)
DT   26-AUG-1993 (first entry)
XX
DE   Sequence associated with myogenic vector systems.
XX
KW   Myogenic vector system; expression; myogenic tissue; vector; vaccine;
KW   insulin-like growth factor; IGF-II; insulin growth factor; milk; meat;
KW   binding protein; growth hormone; growth hormone releasing hormone;
KW   apolipoprotein; A-I; glycogen phosphorylase; alpha-1-antitrypsin; human;
KW   animal; tissue culture; gene replacement; dystrophin; IGF-I;
KW   muscle atrophy; spinal column; injury; neuromuscular disease; MSV;
KW   vitamin D; reulatory element; VDRE; herpes simplex virus; HSV; ss.
XX
OS   Unidentified.
XX
PN   WO9309236-A1.
XX
PD   13-MAY-1993.
XX
PF   03-NOV-1992; 92WO-US009353.
XX
PR   06-NOV-1991; 91US-00789919.
XX
PA   (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI   Schwartz RJ, Demayo FJ, Omalley BW;
XX
DR   WPI; 1993-167701/20.
XX
PT   Myogenic vector systems for expressing nucleic acid sequences - useful in
PT   gene replacement therapy to treat muscle atrophy, atherosclerotic
PT   cardiovascular conditions, etc.
XX
PS   Disclosure; Page 29-30; 56pp; English.
XX
CC   This sequence represents a target sequence which was designed to contain
CC   synthesised multimers of the vitamin D regulatory element (VDRE). This

```
RESULT 2
      AAX88057
ID   AAX88057 standard; cDNA; 27 BP.
AC   AAX88057;
XX
DT   08-SEP-1999  (first entry)
XX
DE   Vitamin D regulatory element (VDRE) cDNA.
XX
KW   Plasmid pIGO100A; plasmid pIG0335; expression vector; treatment; disease;
KW   RNA stability element; gene therapy; muscle atrophy; neurological;
KW   muscular disease; systemic disease; aging; trophic factor; haemophilia;
KW   clotting factor; atherogenesis; atherosclerotic; cardiovascular;
KW   cerebrovascular; peripheral-vascular disease; hormone deficiency;
KW   diabetes; transgenic animal; carcinogen; regulatory element;
KW   livestock improvement; immune response; ss.
OS   Unidentified.
XX
PN   US5925564-A.
XX
PD   20-JUL-1999.
XX
PF   07-JUN-1995;  95US-00472809.
XX
PR   06-NOV-1991;  91US-00789919.
PR   09-MAR-1994;  94US-0020846.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   O'malley BW,  Demayo FJ,  Schwartz RJ;
XX
DR   WPI; 1999-418276/35.
XX
PT   New expression vector system useful for gene therapy.
XX
PS   Disclosure; Col 39-40; 67pp; English.
XX
CC   This invention describes novel expression vector systems containing RNA
CC   stability elements from 3' flanking sequences used for establishing
CC   expression of a nucleic acid sequence within a tissue. The vectors also
CC   facilitate enhanced expression in tissues and target expression with
CC   tissue specificity. The expression vectors can be used to treat diseases
CC   through gene therapy by targeting the vector to specific tissues.
CC   Diseases that can be treated include muscle atrophy associated with
CC   target was linked to the herpes simplex virus (HSV) promoter and used in
CC   a myogenic vector system (MSV) of the invention. The MSV are capable of
CC   expressing any specific nucleic acid sequence in myogenic tissue. This
CC   involves a regulatable myogenic vector system. The vectors can be used
CC   for the expression of eg. insulin-like growth factor (IGF)-1, IGF-II,
CC   insulin growth factor binding protein, growth hormone, growth hormone
CC   releasing hormone, apolipo- protein A-I, glycogen phosphorylase, alpha-1-
CC   antitrypsin or dystrophin. They can be used for the supply of polypeptide
CC   into human, animal or tissue culture, for gene replacement or for vaccine
CC   production. They can be used for eg. treating muscle atrophy in aging
CC   humans or induced by spinal column injuries or neuromuscular diseases,
CC   preventing or treating growth disease. When the vector encode growth
CC   hormone they can be used for increasing milk or meat production in
CC   animals. (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25
CC   -MAR-2003 to correct PN field.)
XX
SQ   Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 U; 0 Other;

Query Match             100.0%;  Score 27;  DB 2;  Length 27;
Best Local Similarity   100.0%;  Pred. No. 0.018;
Matches 27;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 GGTGACTCACCGGGTGAACGGGGGCATT 27
              ||||||||||||||||||||||||||||
Db          1 GGTGACTCACCGGGTGAACGGGGGCATT 27
```

```
RESULT 3
      AAF85650
ID   AAF85650 standard; DNA; 57 BP.
AC   AAF85650;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Human osteocalcin vitamin D response element #2.
XX
KW   Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KW   vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KW   glucocorticoid mediated disorder; granuloma forming disease;
KW   vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FH
FT   misc_feature    14..19
FT                   /*tag= a
FT                   /note= "response element half site"
FT   misc_feature    23..29
FT                   /*tag= b
FT                   /note= "response element half site"
XX
PN   WO200121649-A2.
XX
PD   29-MAR-2001.
XX
PF   20-SEP-2000; 2000WO-US025844.
XX
PR   22-SEP-1999; 99US-00400967.
XX
PA   (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI   Adams JS;
XX
DR   WPI; 2001-308082/32.
XX
PT   New vitamin D response element-binding protein (VDRE-BP) useful in
PT   modifying vitamin D receptor activity in producing anti-VDRE-BP
PT   antibodies, in identifying agonists and antagonists of the protein, or in
PT   gene therapy.
XX
PS   Example; Page 46; 81pp; English.
CC   neurological, muscular or systemic disease, aging by causing tissues to
CC   express trophic factors, haemophilia by causing tissues to express and
CC   secrete clotting factor into the circulation, atherogenesis and
CC   atherosclerotic, cardiovascular, cerebrovascular or peripheral-vascular
CC   disease by causing tissues to express factors involved in tissue
CC   metabolism. They can be used to replace genes of inherited genetic
CC   defects or acquired hormone deficiencies e.g. diabetes. To transform
CC   cells to produce particular proteins or RNA in vitro. To create
CC   transgenic animals which can be used for research into human diseases,
CC   assessing novel therapeutic methods, assessing the effect of chemical and
CC   physical carcinogens and for studying the effect of genes and genetic
CC   regulatory elements or livestock improvement. They can be used to induce
CC   an immune response. These vectors provide controlled expression of the
CC   genes they carry and produce a significantly high level of expression.
CC   Using 3'UTR sequences reduces the decay rates of the mRNAs encoded by the
CC   vectors which causes increased expression
XX
SQ   Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 U; 0 Other;

Query Match             100.0%;  Score 27;  DB 2;  Length 27;
Best Local Similarity   100.0%;  Pred. No. 0.018;
Matches 27;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 GGTGACTCACCGGGTGAACGGGGGCATT 27
              ||||||||||||||||||||||||||||
Db          1 GGTGACTCACCGGGTGAACGGGGGCATT 27
```

```
XX
CC   The present invention provides the protein and coding sequences for two
CC   vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
CC   These confer vitamin D resistance on the primate, a characteristic which
CC   is associated with high circulating levels of other steroid hormones. The
CC   sequences provided by the invention can be used to identify treatments
CC   for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
CC   hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
CC   forming diseases. Note: The present sequence is an example of a vitamin D
CC   response element. Note: The present sequence is stated in the
CC   specification as being the same as that shown as SEQ ID NO: 6 on page 46
CC   (see AAF85639). However, the sequences differ at the last position
XX
SQ   Sequence 57 BP; 12 A; 18 C; 17 G; 10 T; 0 U; 0 Other;

Query Match             92.6%;    Score 25;   DB 4;   Length 57;
Best Local Similarity
Matches   25;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
                                    Pred. No. 0.15;

Qy          1 GGTGACTCACCGGGTGAACGGGGCA 25
              |||||||||||||||||||||||||
Db          3 GGTGACTCACCGGGTGAACGGGGCA 27


RESULT 4
ID   AAV18523 standard; DNA; 53 BP.
XX
AC   AAV18523;
XX
DT   11-JUN-1998   (first entry)
XX
DE   Primer VDRE(OC) for rat vitamin D receptor cDNA.
XX
KW   Rat; vitamin D receptor; isoform protein; VDR1; VDR0; diagnosis;
KW   dominant negative receptor; signal transmission channel;
KW   bone density disorder; screening; PCR primer; ss.
XX
OS   Synthetic.
OS   Rattus rattus.
XX
PN   WO9747172-A1.
XX
PD   18-DEC-1997.
XX
PF   10-JUN-1997;   97WO-IB000947.
XX
PR   10-JUN-1996;   96JP-00194179.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Kato S, Ueno K;
XX
DR   WPI; 1998-051917/05.
XX
PT   DNA encoding a vitamin D receptor isoform protein - useful for bone
PT   density determination and for screening substances for vitamin D
PT   activity.
XX
PS   Example 3; Page 18; 46pp; Japanese.
XX
CC   The present sequence was used in the preparation of the cDNA encoding the
CC   rat vitamin D receptor isoform protein (VDR1). The isoform differs from
CC   the normal receptor (VDR0) in having the vitamin D response element
CC   curtailed by 86 residues, and having an extra 19 residues inserted at the
CC   C-terminal of this element. It acts as a dominant negative receptor in
CC   the vitamin D signal transmission channel. The isoform protein can be
CC   used to diagnose bone density disorders, and screen for substances having
CC   potential vitamin D-like activity
XX
SQ   Sequence 53 BP; 8 A; 13 C; 21 G; 11 T; 0 U; 0 Other;
```

```
Query Match
Best Local Similarity   100.0%;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
Matches   23;                     Pred. No. 1.2;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db         17 GGTGACTCACCGGGTGAACGGGG 39


RESULT 5
ID   AAX16604 standard; DNA; 53 BP.
XX
AC   AAX16604;
XX
DT   29-APR-1999   (first entry)
XX
DE   Human vitamin D receptor isoform protein VDR1 PCR primer OC.
XX
KW   Vitamin D receptor; VDR; bone density; retinoic acid derivative; steroid;
KW   bone calcium regulator; immunosuppressant; anticancer; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9856908-A1.
XX
PD   17-DEC-1998.
XX
PF   13-JUN-1997;   97WO-JP002052.
XX
PR   13-JUN-1997;   97WO-JP002052.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Kato S, Ueno K;
XX
DR   WPI; 1999-080898/07.
XX
PT   Gene for vitamin D receptor isoform protein which blocks vitamin D signal
PT   pathway - and production of recombinant isoform protein using it, for
PT   bone density assay and for screening compounds for vitamin D activity.
XX
PS   Example 3; Page 18; 47pp; Japanese.
XX
CC   The present sequence represents a PCR primer for human vitamin D receptor
CC   (VDR) isoform protein VDR1. VDR1 differs from the canonical form of VDR
CC   (VDR0) by lacking the amino acid sequence encoded by exon 9, but instead
CC   substitutes a short sequence (GTEPGREEELRDLGHVGDCE in the rat protein)
CC   encoded by the 5'-end of intron 8. VDR1 has a dominant negative effect on
CC   the vitamin D signalling pathway. The isoform protein can be used for
CC   determining bone density, and for the screening of compounds (e.g.
CC   steroids and retinoic acid derivatives) for vitamin D activity (e.g. as
CC   bone calcium regulators, immunosuppressants or anticancer agents)
XX
SQ   Sequence 53 BP; 8 A; 13 C; 21 G; 11 T; 0 U; 0 Other;

Query Match
Best Local Similarity   85.2%;    Score 23;   DB 2;   Length 53;
Matches   23;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
                                    Pred. No. 1.2;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db         17 GGTGACTCACCGGGTGAACGGGG 39


RESULT 6
ID   AAF85650 standard; DNA; 57 BP.
XX
AC   AAF85650;
XX
DT   25-JUN-2001   (first entry)
XX
```

Left column

DE   Human osteocalcin vitamin D response element #2.
XX
KW   Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KW   vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KW   glucocorticoid mediated disorder; granuloma forming disease;
KW   vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT
FT   misc_feature   14..19
FT   /*tag= a
FT   /note= "response element half site"
FT   misc_feature   23..29
FT   /*tag= b
FT   /note= "response element half site"
XX
PN   WO20121649-A2.
XX
PD   29-MAR-2001.
XX
PF   20-SEP-2000; 2000WO-US025844.
XX
PR   22-SEP-1999; 99US-00400967.
XX
PA   (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI   Adams JS;
XX
DR   WPI; 2001-308082/32.
XX
PT   New vitamin D response element-binding protein (VDRE-BP) useful in
PT   modifying vitamin D receptor activity, in producing anti-VDRE-BP
PT   antibodies, in identifying agonists and antagonists of the protein, or in
PT   gene therapy.
PS   Example; Page 46; 81pp; English.
XX
CC   The present invention provides the protein and coding sequences for two
CC   vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
CC   These confer vitamin D resistance on the primate, a characteristic which
CC   is associated with high circulating levels of other steroid hormones. The
CC   sequences provided by the invention can be used to identify treatments
CC   for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
CC   hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
CC   forming diseases. The present sequence is an example of a vitamin D
CC   response element. Note: The present sequence is stated in the
CC   specification as being the same as that shown as SEQ ID NO: 6 on page 46
CC   (see AAF85639). However, the sequences differ at the last position
XX
SQ   Sequence 57 BP; 12 A; 18 C; 17 G; 10 T; 0 U; 0 Other;

```
Qy   1 GGTGACTCACCGGGGTGAACGGGG 23
       |||||||||||||||||||||||
Db  55 GGTGACTCACCGGGGTGAACGGGG 33
```

Query Match   85.2%;   Score 23;   DB 4;   Length 57;
Best Local Similarity   100.0%;   Pred. No. 1.2;
Matches 23;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

RESULT 7
ID   AAF85639 standard; DNA; 57 BP.
XX
AC   AAF85639;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Human osteocalcin vitamin D response element.
XX
KW   Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KW   vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KW   glucocorticoid mediated disorder; granuloma forming disease;
KW   vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.

Right column

KW   glucocorticoid mediated disorder; granuloma forming disease;
KW   vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT
FT   misc_feature   14..19
FT   /*tag= a
FT   /note= "response element half site"
FT   misc_feature   23..29
FT   /*tag= b
FT   /note= "response element half site"
XX
PN   WO20121649-A2.
XX
PD   29-MAR-2001.
XX
PF   20-SEP-2000; 2000WO-US025844.
XX
PR   22-SEP-1999; 99US-00400967.
XX
PA   (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI   Adams JS;
XX
DR   WPI; 2001-308082/32.
XX
PT   New vitamin D response element-binding protein (VDRE-BP) useful in
PT   modifying vitamin D receptor activity, in producing anti-VDRE-BP
PT   antibodies, in identifying agonists and antagonists of the protein, or in
PT   gene therapy.
PS   Example; Page 76; 81pp; English.
XX
CC   The present invention provides the protein and coding sequences for two
CC   vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
CC   These confer vitamin D resistance on the primate, a characteristic which
CC   is associated with high circulating levels of other steroid hormones. The
CC   sequences provided by the invention can be used to identify treatments
CC   for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
CC   hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
CC   forming diseases. The present sequence is an example of a vitamin D
CC   response element. Note: The present sequence is stated in the
CC   specification as being the same as that shown as SEQ ID NO: 6 in the
CC   sequence listing (see AAF85650). However, the sequences differ at the
CC   last position
XX
SQ   Sequence 57 BP; 11 A; 18 C; 18 G; 10 T; 0 U; 0 Other;

```
Qy   1 GGTGACTCACCGGGGTGAACGGGG 23
       |||||||||||||||||||||||
Db   3 GGTGACTCACCGGGGTGAACGGGG 25
```

Query Match   85.2%;   Score 23;   DB 4;   Length 57;
Best Local Similarity   100.0%;   Pred. No. 1.2;
Matches 23;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

RESULT 8
ID   AAF85639/c standard; DNA; 57 BP.
XX
AC   AAF85639/c;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Human osteocalcin vitamin D response element.
XX
KW   Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KW   vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KW   glucocorticoid mediated disorder; granuloma forming disease;
KW   vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    14..19
FT                   /*tag= a
FT                   /note= "response element half site"
FT   misc_feature    23..29
FT                   /*tag= b
FT                   /note= "response element half site"
XX
PN   WO20012121649-A2.
XX
PD   29-MAR-2001.
XX
PF   20-SEP-2000; 2000WO-US025844.
XX
PR   22-SEP-1999; 99US-00400967.
XX
PA   (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI   Adams JS;
XX
DR   WPI; 2001-308082/32.
XX
PT   New vitamin D response element-binding protein (VDRE-BP) useful in
PT   modifying vitamin D receptor activity, in producing anti-VDRE-BP
PT   antibodies, in identifying agonists and antagonists of the protein, or in
PT   gene therapy.
XX
PS   Example; Page 76; 81pp; English.
XX
CC   The present invention provides the protein and coding sequences for two
CC   vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
CC   These confer vitamin D resistance on the primate, a characteristic which
CC   is associated with high circulating levels of other steroid hormones. The
CC   sequences provided by the invention can be used to identify treatments
CC   for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
CC   hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
CC   forming diseases. The present sequence is an example of a vitamin D
CC   response element. Note: The present sequence is stated in the
CC   specification as being the same as that shown as SEQ ID NO: 6 in the
CC   sequence listing (see AAF85650). However, the sequences differ at the
CC   last position.
XX
SQ   Sequence 57 BP; 11 A; 18 C; 18 G; 10 T; 0 U; 0 Other;

Query Match        85.2%;  Score 23;  DB 4;  Length 57;
Best Local Similarity   100.0%;  Pred. No. 1.2;
Matches  23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 GGTGACTCACCGGGTGAACGGGG 23
         |||||||||||||||||||||||
Db   56 GGTGACTCACCGGGTGAACGGGG 34


RESULT 9
AAQ79120
ID   AAQ79120 standard; DNA; 301 BP.
XX
AC   AAQ79120;
XX
DT   25-MAR-2003 (revised)
DT   06-AUG-1995 (first entry)
XX
DE   Osteocalcin gene 5' region.
XX
KW   TNFRE; tumor necrosis factor responsive element; osteocalcin; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_RNA        229..237
FT                   /*tag= a
FT                   /label= TNFRE
FT                   /note= "Claim 4, page 42"
XX
PN   WO9428124-A1.
XX
PD   08-DEC-1994.
XX
PF   18-MAY-1994; 94WO-US005659.
XX
PR   20-MAY-1993; 93US-00066372.
XX
PA   (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
PI   Li Y, Stashenko P;
XX
DR   WPI; 1995-022796/03.
XX
PS   Disclosure; Page 32; 63pp; English.
XX
CC   Tumor necrosis factor (TNF)-mediated down-regulation of the osteocalcin
CC   gene is exerted through a TNF responsive element (TNFRE). Deletion
CC   analysis localized the TNFRE to -523/-515 of the 5' flanking region of
CC   the osteocalcin gene. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 301 BP; 41 A; 111 C; 84 G; 65 T; 0 U; 0 Other;

Query Match        85.2%;  Score 23;  DB 2;  Length 301;
Best Local Similarity   100.0%;  Pred. No. 1.4;
Matches  23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 GGTGACTCACCGGGTGAACGGGG 23
         |||||||||||||||||||||||
Db  240 GGTGACTCACCGGGTGAACGGGG 262


RESULT 10
AB258131
ID   AB258131 standard; DNA; 565 BP.
XX
AC   AB258131;
XX
DT   22-APR-2003 (first entry)
XX
DE   Human osteocalcin promoter.
XX
KW   Human; osteocalcin; promoter; bone; tumour; prostate cancer; metastasis;
KW   gene therapy; diagnosis; prognosis; marker; cytostatic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_signal     74..91
FT                   /*tag= a
FT                   /note= "vitamin D response element"
FT   misc_signal     414..431
FT                   /*tag= c
FT                   /note= "OSE2 element"
FT   misc_signal     515..543
FT                   /*tag= b
FT                   /note= "OSE1 element"
XX
PN   WO2003006621-A2.
XX
PD   23-JAN-2003.
XX
PF   12-JUL-2002; 2002WO-US022216.
XX
PR   13-JUL-2001; 2001US-0305360P.
XX
```

PA (UIVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Chung LWK, Yeung F;
XX
DR WPI; 2003-221733/21.
XX
PT Nucleic acid sequence for diagnosing, prognosing or treating calcified
PT tumors and tissue, e.g. prostate cancer, comprises an improved
PT recombinant human osteocalcin promoter activity.
XX
PS Claim 12; Page 51-52; 55pp; English.
XX
CC The present sequence is that of the human osteocalcin (hOC) promoter. An
CC evaluation of the hOC promoter was conducted in which the functional
CC hierarchy of the cis-acting elements OSE1, OSE2 and AP-1/VDRE (vitamin D
CC response element) was defined in an androgen-independent human prostate
CC cancer PC-3 cell line. By juxtaposing dimers of these 3 elements, a
CC minimal hOC super-promoter (see AB258130) was produced. The super-
CC promoter displays over 8-fold higher activity than the native full-length
CC hOC promoter in a tissue-specific manner in PC-3 cells. Expression
CC vectors comprising the hOC promoter sequences and transformed host cells
CC are provided by the invention. In one embodiment, the hOC super-promoter
CC is operably linked to a nucleic acid encoding a heterologous protein,
CC ribozyme, dominant-negative or antisense RNA and used to deliver
CC therapeutic genes to localised or disseminated tumours. The hOC promoter
CC sequences can also be used to deliver therapeutic genes to fractured
CC bones for bone repair. hOC promoter activation by extracellular matrices
CC and soluble factors secreted by prostate cancer and bone cells in useful
CC as a marker for the diagnosis and prognosis of prostate cancer
XX
SQ Sequence 565 BP; 101 A; 185 C; 170 G; 109 T; 0 U; 0 Other;

Query Match   85.2%; Score 23; DB 8; Length 565;
Best Local Similarity   100.0%; Pred. No. 1.5;
Matches 23;   Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db   74 GGTGACTCACCGGGGTGAACGGGG 96

RESULT 11

ID AAT90018 standard; DNA; 597 BP.
XX
AC AAT90018;
XX
DT 20-NOV-1997 (first entry)
XX
DE Bases -575 to 22 of the human osteocalcin gene.
XX
KW Recombinant vector; human; osteocalcin; promoter; transcription;
KW control region; determination; activity; vitamin D derivative;
KW bone metabolism; regulation; ss.
XX
OS Homo sapiens.
XX
PN JP0182586-A.
XX
PD 15-JUL-1997.
XX
PF 29-DEC-1995; 95JP-0035281?.
PR 29-DEC-1995; 95JP-0035281?.
XX
PA (SRLS-) SRL KK.
XX
DR WPI; 1997-409886/38.
XX
PT Recombinant vector comprising vitamin D response element - and human
PT osteocalcin gene promoter sequence and transcriptional control region,
PT upstream of structural gene encoding detectable compound.
XX

PS Claim 3; Page 5; 7pp; Japanese.
XX
CC A novel recombinant vector, comprises a vitamin D response element
CC (VDRE), and a human osteocalcin gene promoter sequence and
CC transcriptional control region, i.e. the present sequence, upstream of a
CC structural gene encoding a detectable compound. The vector can be used to
CC determine the activity of vitamin D derivatives as bone metabolism
CC regulators, without using animals. A transformant containing the vector
CC is incubated on a dish to confluence, and a vitamin D derivative added at
CC a concentration of 10-11 to 10-8 M. The cells are incubated for 2 days
CC and recovered, and the yield of the detectable compound determined. When
CC the vitamin D derivative is highly active the yield of the detectable
CC compound is increased, as the osteocalcin gene promoter is activated
XX
SQ Sequence 597 BP; 107 A; 190 C; 187 G; 113 T; 0 U; 0 Other;

Query Match   85.2%; Score 23; DB 2; Length 597;
Best Local Similarity   100.0%; Pred. No. 1.5;
Matches 23;   Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db   65 GGTGACTCACCGGGGTGAACGGGG 87

RESULT 12

ID AAD14384 standard; DNA; 597 BP.
XX
AC AAD14384;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human osteocalcin promoter DNA.
XX
KW Human; antiinflammatory; recombinant human interleukin-4; rhIL-4; OPC;
KW osteoprogenitor cell; odontoprogenitor cell; alveolar bone loss; RA;
KW rheumatoid arthritis; periodontitis; osteolysis; bone metastasis;
KW osteopathic; cytostatic; anticaries; gene-therapy; osteoblast;
KW osteocalcin; dB.
XX
OS Homo sapiens.
XX
PN WO200160983-A2.
XX
PD 23-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US004274.
PR 18-FEB-2000; 2000US-0050739.
XX
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
PI Keeping HS, Reichner JS;
XX
DR WPI; 2001-514769/56.
XX
PT An isolated odontoprogenitor cell or osteoprogenitor cell for the
PT treatment of rheumatoid arthritis both comprising a nucleic acid encoding
PT an anti-inflammatory peptide.
XX
PS Example 3; Page 4; 28pp; English.
XX
CC The patent discloses compositions and methods to deliver an anti-
CC inflammatory composition, e.g. recombinant human interleukin-4 (rhIL-4),
CC to build (or rebuild) bone tissue. The patent also relates to
CC osteoprogenitor cells (OPCs) or odontoprogenitor cells comprising a
CC nucleic acid encoding an anti-inflammatory peptide. The odonto-
CC progenitor cells are useful for treating a mammal at risk of developing
CC periodontitis or is suffering from or is at risk of developing alveolar
CC bone loss due to periodontal disease. OPCs are useful for treating a
CC mammal suffering from or at risk of developing rheumatoid arthritis (RA),
CC periapical or endochondral bone loss, artificial joint particle- induced

CC osteolysis, or osteolytic bone metastases. They are used to inhibit the
CC formation of osteoclasts which resorb bone and to stimulate the growth of
CC new bones. The OPCs and odontoprogenitor cells are useful for inhibiting
CC osteolysis and for inducing differentiation of a bone marrow stromal
CC cells. These cells are also used in gene therapy. The present sequence is
CC human osteoblast-specific promoter DNA, osteocalcin promoter
XX
SQ Sequence 597 BP; 107 A; 190 C; 187 G; 113 T; 0 U; 0 Other;

Query Match   85.2%; Score 23; DB 5; Length 597;
Best Local Similarity   100.0%;   Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGCGGGTGAACGGGG 23
        |||||||||||||||||||||||
Db 65 GGTGACTCACCGCGGGTGAACGGGG 87

RESULT 13
ID   AAD48137 standard; DNA; 2650 BP.
XX
AC   AAD48137;
XX
DT   24-FEB-2003 (first entry)
XX
DE   Human pleiotrophin/OSF-1 DNA.
XX
KW   Brain; tumour protein target; Tbt; ischaemic stroke; cancer; epilepsy;
KW   schizophrenia; depression; Alzheimer's disease; Parkinson's disease;
KW   Huntington's chorea; traumatic head injury; dementia; stupor; headache;
KW   coma; vertigo; weakness; myasthenia gravis; cerebrovascular disorder;
KW   infection; multiple sclerosis; pregnancy; medical illness; vasotropic;
KW   metabolic deficiency; cerebroprotective; antidepressant; antibacterial;
KW   cytostatic; nootropic; analgesic; fungicide; virucide; pleiotrophin;
KW   OSF-1; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1542..2048
FT   /*tag= a
FT   /product= "Human pleiotrophin/OSF-1 protein"
XX
PN   WO0276510-A1.
XX
PD   03-OCT-2002.
XX
PF   22-MAR-2002; 2002WO-US008992.
XX
PR   23-MAR-2001; 2001US-00816703.
PR   17-OCT-2001; 2001US-00983000.
XX
PA   (AGYT-) AGY THERAPEUTICS INC.
XX
PI   Mueller S, Melcher T, Chin DJ;
XX
DR   WPI; 2003-029903/02.
DR   P-PSDB; AAE30342.
XX
PT   Developing active agents that modulate the activity of a brain tumor
PT   protein target gene or gene product for treating the brain tumour e.g.
PT   comprises contacting an agent with a brain tumor protein.
XX
PS   Claim 1; Page 130-132; 135pp; English.
XX
CC   The invention relates to a method for developing biologically active
CC   agents that modulate activity of a brain tumour protein target (Tbt) gene
CC   or gene product. The method is useful for developing biologically active
CC   agents that modulate the activity of a brain tumour protein target gene
CC   or gene product. Compounds that bind to the brain tumour proteins are
CC   useful for treating e.g. ischaemic stroke, brain cancer, epilepsy,
CC   schizophrenia, depression, Alzheimer's disease, Parkinson's disease,
CC   Huntington's chorea, traumatic head injury, dementia, stupor, headache,
CC   coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders,
CC   infectious disorders (including fungal, bacterial, viral and parasitic
CC   infections), multiple sclerosis, and other complications associated with
CC   pregnancy, medical illness, alcohol and substance abuse, toxins and
CC   metabolic deficiencies. The brain tumour proteins may also be used to
CC   raise antibodies. The present sequence is human pleiotrophin/ OSF-1 DNA
CC   used to illustrate the method of the invention
XX
SQ   Sequence 2650 BP; 639 A; 657 C; 649 G; 705 T; 0 U; 0 Other;

Query Match   85.2%; Score 23; DB 8; Length 2650;
Best Local Similarity   100.0%;   Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGCGGGTGAACGGGG 23
        |||||||||||||||||||||||
Db 343 GGTGACTCACCGCGGGTGAACGGGG 365

RESULT 14
ID   ADN06011
ID   ADN06011 standard; cDNA; 2650 BP.
XX
AC   ADN06011;
XX
DT   01-JUL-2004 (first entry)
XX
DE   Antipsoriatic cDNA sequence #1241.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   P-PSDB; ADN06012.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 1; SEQ ID NO 2406; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polynucleotides of the invention.
XX
SQ   Sequence 2650 BP; 657 A; 649 G; 705 T; 0 U; 0 Other;

Query Match   85.2%; Score 23; DB 12; Length 2650;
Best Local Similarity   100.0%;   Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGCGGGTGAACGGGG 23
        |||||||||||||||||||||||
Db 343 GGTGACTCACCGCGGGTGAACGGGG 365

```
RESULT 15
ABZ34742
ID   ABZ34742 standard; cDNA; 2657 BP.
XX
AC   ABZ34742;
XX
DT   04-FEB-2003  (first entry)
XX
DE   Coding sequence SEQ ID 100, downregulated in osteogenesis.
XX
KW   Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW   osteoporosis; bone disease; downregulator; human; ss.
XX
OS   Homo sapiens.
XX
XX
PN   WO200281745-A2.
XX
PD   17-OCT-2002.
XX
PF   05-APR-2002; 2002WO-IB002211.
XX
PR   05-APR-2001; 2001US-0281400P.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
PI   Connolly T, Jackson A, Bushnell SE, Rawadi G;
XX
DR   WPI; 2003-058567/05.
XX
PS   Claim 27; Page 126-127; 237pp; English.
XX
PT   Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT   useful for bone disease therapy in subject.
XX
CC   The present invention relates to novel nucleotide sequences, which are
CC   differentially expressed in models of osteogenesis upon being put in
CC   contact with a stimulator of osteogenesis. The present sequence is one
CC   such sequence. This sequence can be used for diagnosing osteoporosis/bone
CC   disease in a patient, promoting osteogenesis and/or preventing
CC   osteoporosis/bone disease. The present sequence encodes a secreted
CC   protein
XX
SQ   Sequence 2657 BP; 640 A; 656 C; 648 G; 706 T; 0 U; 7 Other;

Query Match             85.2%;  Score 23;  DB 8;  Length 2657;
Best Local Similarity   100.0%;  Pred. No. 1.7;
Matches   23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GGTGACTCACCGGGGTGAACGGGG 23
              ||||||||||||||||||||||||
Db        343 GGTGACTCACCGGGGTGAACGGGG 365
```

Search completed: November 29, 2004, 14:21:32
Job time : 403 secs

Title: 27
Perfect score: 27
Sequence: 1 GGTGACTCACCGGGGTGAACGGGGCATT 27

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:01 ; Search time 89 seconds
                 GenCore version 5.1.6
                 Copyright (c) 1993 - 2004 Compugen Ltd.
                 (without alignments)
                 215.633 Million cell updates/sec

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	100.0	27	27	1	US-07-789-919A-2	Sequence 2, Appli
2	100.0	27	27	1	US-08-209-846A-2	Sequence 2, Appli
3	100.0	27	27	3	US-08-472-809B-2	Sequence 2, Appli
4	100.0	27	27	3	US-08-438-265-2	Sequence 2, Appli
5	85.2	23	301	5	PCT-US94-05659-2	Sequence 1, Appl
6	67.4	18.2	1440	4	US-10-141-634-1	Sequence 1, Appli
7	66.7	18	624	4	US-09-489-039A-1953	Sequence 1953, Ap
8	66.7	18	768	4	US-09-489-039A-1930	Sequence 1930, Ap
9	66.7	18	4245	4	US-09-276-531-16	Sequence 16, Appl
10	66.7	18	4500	4	US-09-087-031E-27	Sequence 27, Appl
11	66.7	18	4411529	3	US-09-103-840A-2	Sequence 2, Appli
12	66.7	18	4411529	3	US-09-103-840A-1	Sequence 1, Appli
13	65.2	17.6	2138	4	US-09-056-556-183	Sequence 183, App
14	65.2	17.6	2138	4	US-09-072-599-178	Sequence 178, App
15	65.2	17.6	2138	4	US-09-072-967-183	Sequence 183, App
16	64.4	17.4	7194	4	US-09-601-326-76	Sequence 76, Appl
17	64.4	17.4	15420	4	US-09-601-326-54	Sequence 54, Appl
18	63.0	17	790	4	US-09-976-594-796	Sequence 796, App
19	63.0	17	2933	3	US-09-262-773-209	Sequence 209, App
20	63.0	17	23071	3	US-09-262-773-210	Sequence 210, App
21	63.0	17	4403765	3	US-09-103-840A-2	Sequence 2, Appli
22	62.2	16.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
23	62.2	16.8	5917	4	US-09-780-175-17	Sequence 17, Appl
24	61.5	16.6	321	4	US-09-401-064-348	Sequence 348, App
25	61.5	16.6	14462	4	US-09-84-250-9	Sequence 9, Appli
26	60.7	16.4	30	4	US-08-865-960-27	Sequence 27, Appl
27	60.7	16.4	629	4	US-08-477-135A-123	Sequence 123, App
28	60.7	16.4	630	4	US-09-489-039A-938	Sequence 938, App
29	60.7	16.4	1170	4	US-09-489-039A-936	Sequence 936, App
30	60.7	16.4	1629	4	US-09-489-039A-4586	Sequence 4586, Ap
31	60.0	16.2	3177	4	US-09-252-991A-12534	Sequence 12534, A
32	59.3	16	192	2	US-08-604-999A-6	Sequence 6, Appli
33	59.3	16	212	4	US-09-016-434-50	Sequence 50, Appl
34	59.3	16	243	4	US-09-489-039A-5679	Sequence 5679, Ap
35	59.3	16	259	4	US-09-513-999C-2402	Sequence 2402, Ap
36	59.3	16	295	4	US-09-489-172-44	Sequence 44, Appl
37	59.3	16	411	3	US-09-059-369-16	Sequence 16, Appl
38	59.3	16	596	4	US-09-583-110-2085	Sequence 2085, Ap
39	59.3	16	735	4	US-09-222-575-67	Sequence 67, Appl
40	59.3	16	1022	4	US-09-389-681-67	Sequence 67, Appl
41	59.3	16	1022	4	US-09-620-405B-67	Sequence 67, Appl
42	59.3	16	1022	4	US-09-339-338-67	Sequence 67, Appl
43	59.3	16	1022	4	US-09-433-826B-67	Sequence 67, Appl
44	59.3	16	1022	4	US-09-604-287A-67	Sequence 67, Appl
45	59.3	16	1022	4	US-09-604-287A-67	Sequence 67, Appl

RESULT 1
US-07-789-919A-2
Sequence 2, Application US/07789919A
Patent No. 5298422
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Myogenic Vector Systems
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas D. Paul
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/789,919A
FILING DATE: 19911106
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5379
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-789-919A-2

ALIGNMENTS

Query Match   100.0%; Score 27; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGGCATT 27

Db
1 GGTGACTCACCGGGTGAACGGGGCATT 27
||||||||||||||||||||||||||||

RESULT 2
US-08-209-846A-2
Sequence 2, Application US/08209846A
Patent No. 5756264
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Expression Vector Systems and Methods of Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth St.
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,846A
FILING DATE: 09-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/789,919
FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Knight, Matthew W.
REGISTRATION NUMBER: 36,846
REFERENCE/DOCKET NUMBER: 204/302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-489-1600
TELEFAX: 213-955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-209-846A-2

Qy
1 GGTGACTCACCGGGTGAACGGGGCATT 27
||||||||||||||||||||||||||||
Db
1 GGTGACTCACCGGGTGAACGGGGCATT 27

Query Match   100.0%; Score 27; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-472-809B-2

RESULT 3
US-08-472-809B-2
Sequence 2, Application US/08472809B
Patent No. 5925564
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco J.
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Expression Vector Systems and Method of Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,809B
FILING DATE: June 7, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,846
FILING DATE: March 9, 1994
APPLICATION NUMBER: 07/789,919
FILING DATE: No. 5925564ember 6, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 214/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-472-809B-2

Qy
1 GGTGACTCACCGGGTGAACGGGGCATT 27
||||||||||||||||||||||||||||
Db
1 GGTGACTCACCGGGTGAACGGGGCATT 27

Query Match   100.0%; Score 27; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-438-265-2

RESULT 4
US-08-438-265-2
Sequence 2, Application US/08438265
Patent No. 6361984
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Expression Vector Systems and Methods of Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth St.
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,265
FILING DATE: 09-May-1995

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/368,776
  FILING DATE: <Unknown>
  APPLICATION NUMBER: US 07/789,919
  FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Knight, Matthew W.
  REGISTRATION NUMBER: 36,846
  REFERENCE/DOCKET NUMBER: 204/302
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 213-489-1600
  TELEFAX: 213-955-0440
  TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 27 bases
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-438-265-2

Query Match   100.0%;   Score 27;   DB 3;   Length 27;
Best Local Similarity   100.0%;   Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 GGTGACTCACCGGGTGAACGGGGCATT 27
        |||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGTGAACGGGGCATT 27
```

RESULT 5
PCT-US94-05659-2
Sequence 2, Application PC/TUS9405659
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION: TNF` RESPONSIVE ELEMENT, TNF`-INDUCED DNA-BINDING
  NUMBER OF SEQUENCES: 24
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
    STREET: Two Militia Drive
    CITY: Lexington
    STATE: Massachusetts
    COUNTRY: U.S.A.
    ZIP: 02173
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: PCT/US94/05659
    FILING DATE:
    CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Granahan, Patricia
    REGISTRATION NUMBER: 32,227
    REFERENCE/DOCKET NUMBER: FDC93-01 FF
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 617-861-6240
    TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 301 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)

Query Match   85.2%;   Score 23;   DB 5;   Length 301;

Best Local Similarity   100.0%;   Pred. No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 GGTGACTCACCGGGTGAACGGGG 23
        ||||||||||||||||||||||
Db 240 GGTGACTCACCGGGTGAACGGGG 262
```

RESULT 6
US-10-141-634-1/c
Sequence 1, Application US/10141634
Patent No. 6734010
GENERAL INFORMATION:
  APPLICANT: Yu, Xuanchuan
  APPLICANT: Xie, Qiongshu
  APPLICANT: Abuin, Alejandro
  APPLICANT: Walke, D. Wade
  TITLE OF INVENTION: No. 6734010el Human Kinases and Polynucleotides Encoding the Sar
  FILE REFERENCE: LEX-0344-USA
  CURRENT APPLICATION NUMBER: US/10/141,634
  CURRENT FILING DATE: 2002-05-08
  PRIOR APPLICATION NUMBER: US 60/289,727
  PRIOR FILING DATE: 2001-05-09
  NUMBER OF SEQ ID NOS: 4
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 1
  LENGTH: 1440
  TYPE: DNA
  ORGANISM: homo sapiens
US-10-141-634-1

Query Match   67.4%;   Score 18.2;   DB 4;   Length 1440;
Best Local Similarity   87.0%;   Pred. No. 24;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   5 ACTCACCGGGTGAACGGGGCATT 27
        ||||||| ||||||||||||||
Db 812 ACCCTCGGGGTGAACGGGGCATT 790
```

RESULT 7
US-09-489-039A-1953
Sequence 1953, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
  APPLICANT: Gary Breton et. al
  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
  TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
  FILE REFERENCE: 2709.2004001
  CURRENT APPLICATION NUMBER: US/09/489,039A
  CURRENT FILING DATE: 2000-01-27
  PRIOR APPLICATION NUMBER: US 60/117,747
  PRIOR FILING DATE: 1999-01-29
  NUMBER OF SEQ ID NOS: 14342
  SEQ ID NO 1953
  LENGTH: 624
  TYPE: DNA
  ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1953

Query Match   66.7%;   Score 18;   DB 4;   Length 624;
Best Local Similarity   80.8%;   Pred. No. 26;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy   2 GTGACTCACCGGGTGACGGGGCATT 27
        ||||||||||||| ||| ||||||
Db 434 GTGAGTCACCGGGATGAACGCGGACTT 459
```

RESULT 8
US-09-489-039A-1930/c
Sequence 1930, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
; APPLICANT: Gary Breton et, al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR APPLICATION FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1930
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae

Qy Query Match 66.7%; Score 18; DB 4; Length 768;
Best Local Similarity 80.8%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy     2 GTGACTCACCGGGGTGAACGGGGGCATT 27
           |||||||||| ||||| |||||||||||
Db   291 GTGAGTCACCGGAGTGAAACGCGGACTT 266
```

US-09-489-039A-1930

RESULT 9
US-09-276-531-16/c
; Sequence 16, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT21
; CLONE: 2526432
US-09-276-531-16

Qy Query Match 66.7%; Score 18; DB 3; Length 4245;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      2 GTGACTCACCGGGGTGAACGGGGGCATT 27
Db   1045 GTGACCCACCGGGTTCCCGGGGGCACT 1020
```

RESULT 10
US-09-087-031E-27/c
; Sequence 27, Application US/09087031E
; Patent No. 6479255
; GENERAL INFORMATION:
; APPLICANT: Rubin, Jeffrey S.
; APPLICANT: Finch, Paul
; APPLICANT: Aaronson, Stuart
; APPLICANT: He, Xi
; TITLE OF INVENTION: HUMAN FRP AND FRAGMENTS THEREOF INCLUDING METHODS FOR USING THEM
; FILE REFERENCE: 11613.13US11
; CURRENT APPLICATION NUMBER: US/09/087,031E
; CURRENT FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/087,031
; PRIOR APPLICATION FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/050,417
; PRIOR APPLICATION FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 4500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-087-031E-27

Qy Query Match 66.7%; Score 18; DB 4; Length 4500;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      2 GTGACTCACCGGGGTGAACGGGGGCATT 27
Db   1359 GTGAGCCACCGGGGTTCCCGGGGGGCACT 1334
```

RESULT 11
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEICHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

RESULT 12
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
;　APPLICANT: FLEISCHMANN, Robert D.
;　APPLICANT: WHITE, Owen R.
;　APPLICANT: FRASER, Claire M.
;　APPLICANT: VENTER, John C.
;　TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;　TITLE OF INVENTION: TUBERCULOSIS
;　FILE REFERENCE: 24366-20007.00
;　CURRENT APPLICATION NUMBER: US/09/103,840A
;　CURRENT FILING DATE: 1998-06-24
;　NUMBER OF SEQ ID NOS: 2
;　SOFTWARE: Patentin Ver. 2.1
;　SEQ ID NO 1
;　LENGTH: 4411529
;　TYPE: DNA
;　ORGANISM: Mycobacterium tuberculosis
;　OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match　　66.7%;　Score 18;　DB 3;　Length 4403765;
Best Local Similarity　80.8%;　Pred. No. 63;
Matches　21;　Conservative　0;　Mismatches　5;　Indels　0;　Gaps　0;

Qy　　　2 GTGACTCACCGGGTGAACGGGGCATT 27
　　　　　|||||||||||| ||||||| ||||
Db　2204211 GTGTCGCACCCCGGTGACCGGCGCATT 2204186

Query Match　　66.7%;　Score 18;　DB 3;　Length 4411529;
Best Local Similarity　80.8%;　Pred. No. 63;
Matches　21;　Conservative　0;　Mismatches　5;　Indels　0;　Gaps　0;
US-09-103-840A-1

Qy　　　2 GTGACTCACCGGGTGAACGGGGCATT 27
　　　　　|||||||||||| ||||||| ||||
Db　2206912 GTGTCGCACCCGGTGACCGGCGCATT 2206887

RESULT 13
US-09-056-556-183
Sequence 183, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
;　APPLICANT: Reed, Steven G.
;　APPLICANT: Skeiky, Yasir A.W.
;　APPLICANT: Dillon, Davin C.
;　TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND TREATM
;　NUMBER OF SEQUENCES: 241
;　CORRESPONDENCE ADDRESS:
;　ADDRESSEE: SEED and BERRY LLP
;　STREET: 6300 Columbia Center, 701 Fifth Avenue
;　CITY: Seattle
;　STATE: Washington
;　COUNTRY: USA
;　ZIP: 98104-7092
;　COMPUTER READABLE FORM:
;　MEDIUM TYPE: Floppy disk
;　COMPUTER: IBM PC compatible
;　OPERATING SYSTEM: PC-DOS/MS-DOS
;　SOFTWARE: Patentin Release #1.0, Version #1.30
;　CURRENT APPLICATION DATA:
;　APPLICATION NUMBER: US/09/056,556
;　FILING DATE: 07-APR-1998
;　CLASSIFICATION:
;　ATTORNEY/AGENT INFORMATION:
;　NAME: Maki, David J.
;　REGISTRATION NUMBER: 31,392
;　REFERENCE/DOCKET NUMBER: 210121.457
;　TELECOMMUNICATION INFORMATION:

RESULT 14
US-09-072-596-178
Sequence 178, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
;　APPLICANT: Reed, Steven G.
;　APPLICANT: Skeiky, Yasir A.W.
;　APPLICANT: Dillon, Davin C.
;　APPLICANT: Campos-Neto, Antonia
;　APPLICANT: Houghton, Raymond
;　APPLICANT: Vedvick, Thomas S.
;　APPLICANT: Twardzik, Daniel R.
;　APPLICANT: Lodes, Michael J.
;　APPLICANT: Hendrickson, Ronald C.
;　TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
;　NUMBER OF SEQUENCES: 350
;　CORRESPONDENCE ADDRESS:
;　ADDRESSEE: SEED and BERRY LLP
;　STREET: 6300 Columbia Center, 701 Fifth Avenue
;　CITY: Seattle
;　STATE: Washington
;　COUNTRY: USA
;　ZIP: 98104-7092
;　COMPUTER READABLE FORM:
;　MEDIUM TYPE: Floppy disk
;　COMPUTER: IBM PC compatible
;　OPERATING SYSTEM: PC-DOS/MS-DOS
;　SOFTWARE: Patentin Release #1.0, Version #1.30
;　CURRENT APPLICATION DATA:
;　APPLICATION NUMBER: US/09/072,596
;　FILING DATE: 05-MAY-1998
;　CLASSIFICATION:
;　ATTORNEY/AGENT INFORMATION:
;　NAME: Maki, David J.
;　REGISTRATION NUMBER: 31,392
;　REFERENCE/DOCKET NUMBER: 210121.417C9
;　TELECOMMUNICATION INFORMATION:
;　TELEPHONE: (206) 622-4900
;　TELEFAX: (206) 682-6031
;　INFORMATION FOR SEQ ID NO: 178:
;　SEQUENCE CHARACTERISTICS:
;　LENGTH: 2138 base pairs
;　TYPE: nucleic acid
;　STRANDEDNESS: single
;　TOPOLOGY: linear
US-09-072-596-178

Query Match　　65.2%;　Score 17.6;　DB 4;　Length 2138;
Best Local Similarity　83.3%;　Pred. No. 47;
Matches　20;　Conservative　0;　Mismatches　4;　Indels　0;　Gaps　0;

Qy　　　1 GGTGACTCACCGGGTGAACGGGGGC 24
　　　　　|||| ||||| ||||| ||| | |||
Db　1798 GGTGGCTCAGCGGGTCAACGAGGC 1821

Query Match　　65.2%;　Score 17.6;　DB 3;　Length 2138;
Best Local Similarity　83.3%;　Pred. No. 47;
Matches　20;　Conservative　0;　Mismatches　4;　Indels　0;　Gaps　0;
US-09-056-556-183

Qy　　　1 GGTGACTCACCGGGTGAACGGGGC 24
　　　　　|||| ||||| ||||| ||| |||
Db　1798 GGTGGCTCAGCGGGTCAACGAGGC 1821

RESULT 15
US-09-072-967-183
; Sequence 183, Application US/09072967
; Patent No. 6528877
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 183:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-072-967-183

Query Match   65.2%; Score 17.6; DB 4; Length 2138;
Best Local Similarity   83.3%; Pred. No. 47;
Matches   20; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 GGTGACTCACCGGTGAACGGGGGC 24
              |||| ||||||| |||| ||||||
Db   1798 GGTGGCTCAGCGGGTCAACGAGGC 1821

Search completed: November 29, 2004, 13:46:57
Job time : 103 secs

```
OM nucleic - nucleic search, using sw model

Run on:        November 29, 2004, 14:14:46 ; Search time 4972 seconds
                                              (without alignments)
                                              29.430 Million cell updates/sec

                         Copyright (c) 1993 - 2004 Compugen Ltd.
                         GenCore version 5.1.6

Title:         US-10-073-064-2
Perfect score: 27
Sequence:      1 GGTGACTCACCGGGTGAACGGGGGCATT 27

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      3640242 seqs, 2709731945 residues

Total number of hits satisfying chosen parameters:      7280484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Published Applications NA:*
           1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
           2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
           3:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
           4:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
           5:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
           6:  /cgn2_6/ptodata/2/pubpna/US103_PUBCOMB.seq:*
           7:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
           8:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
           9:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
          10:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
          11:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
          12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
          13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
          14:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
          15:  /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
          16:  /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
          17:  /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
          18:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
          19:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
          20:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
          21:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
```

```
                                    SUMMARIES

Result                 %    Query
No.    Score  Match  Length  DB ID                     Description
------ -----  -----  ------  -----                     -----------
  1     27   100.0    27    US-10-073-064-2           Sequence 2, Appli
  2     23    85.2    597   US-10-683-516-3           Sequence 3, Appli
  3     23    85.2   2657   US-10-171-311-193         Sequence 193, App
  4     18.2  67.4    132   US-10-029-386-19439       Sequence 19439, A
  5     18.2  67.4    391   US-10-027-632-26905       Sequence 26905, A
  6     18.2  67.4    391   US-10-027-632-26905       Sequence 26905, A
  7     18.2  67.4    545   US-10-029-386-5683        Sequence 5683, Ap
  8     18.2  67.4    774   US-10-165-800-23          Sequence 23, Appl
  9     18.2  67.4    995   US-10-311-034-51          Sequence 51, Appl
 10     18.2  67.4   1223   US-09-764-868-257         Sequence 257, App
 11     18.2  67.4   1440   US-10-141-654-1           Sequence 1, Appl
 12     18.2  67.4   1452   US-10-165-800-21          Sequence 21, Appl
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
  13    18.2  67.4   1579   US-10-094-749-1381        Sequence 1381, Ap
  14    18.2  67.4   1581   US-10-478-146-17          Sequence 17, Appl
  15    18.2  67.4   1619   US-10-104-047-1657        Sequence 1657, Ap
  16    18.2  67.4   1778   US-10-098-841-257         Sequence 257, App
  17    18.2  67.4   5340   US-10-311-455-207         Sequence 207, App
  18    18    66.7   1063   US-10-282-122A-22929      Sequence 22929, A
  19    18    66.7   2602   US-10-666-851-1           Sequence 1, Appli
  20    18    66.7   3215   US-09-925-301-51          Sequence 51, Appl
  21    18    66.7   4462   US-09-974-298-28          Sequence 28, Appl
  22    18    66.7   4500   US-10-138-434A-27         Sequence 27, Appl
  23    18    66.7  21230   US-10-087-192-601         Sequence 601, App
  24    17.6  65.2   1075   US-10-425-114-1745        Sequence 1745, Ap
  25    17.6  65.2   1102   US-10-425-114-16946       Sequence 16946, A
  26    17.6  65.2   1518   US-10-739-930-3057        Sequence 3057, Ap
  27    17.6  65.2   2138   US-10-193-002-178         Sequence 178, App
  28    17.6  65.2   2138   US-10-084-843-183         Sequence 183, App
  29    17.6  65.2   2967   US-09-938-842A-432        Sequence 432, App
  30    17.6  65.2   2967   US-09-938-842A-432        Sequence 432, App
  31    17.6  65.2   3323   US-10-225-066A-151        Sequence 151, App
  32    17.6  65.2   3323   US-10-374-780A-237        Sequence 237, App
  33    17.6  65.2  16181   US-10-764-847-1426        Sequence 1426, Ap
  34    17.6  65.2  16181   US-09-764-891-6956        Sequence 6956, Ap
  35    17.6  65.2  16181   US-09-092-154-1426        Sequence 1426, Ap
  36    17.6  65.2 495269   US-10-398-221-8           Sequence 8, Appli
  37    17.6  65.2 3011208   US-10-398-221-2058       Sequence 2058, Ap
  38    17.4  64.4    686   US-11-424-599-110292      Sequence 110292, A
  39    17.4  64.4    864   US-10-369-493-24087       Sequence 24087, A
  40    17.4  64.4    947   US-10-027-632-161355      Sequence 161355,
  41    17.4  64.4    947   US-10-027-632-161355      Sequence 161355,
  42    17.4  64.4    947   US-10-027-632-161356      Sequence 161356,
  43    17.4  64.4   1587   US-10-260-046-15          Sequence 15, Appl
  44    17.4  64.4   1587   US-10-260-046-7           Sequence 7, Appli
  45    17.4  64.4   1611   US-10-260-046-7
```

```
                                    ALIGNMENTS

RESULT 1
US-10-073-064-2
Sequence 2, Application US/10073064
Publication No. US20030104407A1

GENERAL INFORMATION:

APPLICANT: Schwartz, Robert J.
           DeMayo, Franco
           O'Malley, Bert W.

TITLE OF INVENTION: Expression Vector Systems and Methods of
                    Use

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
    ADDRESSEE: Lyon & Lyon
    STREET: 611 West Sixth St.
    CITY: Los Angeles
    STATE: CA
    COUNTRY: USA
    ZIP: 90017

COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/10/073,064
    FILING DATE: 12-Feb-2002
    CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/08/438,265
    FILING DATE: 09-May-1995
    APPLICATION NUMBER: 08/368,776
    FILING DATE: <Unknown>
    APPLICATION NUMBER: US 07/789,919
    FILING DATE: 06-NOV-1991

ATTORNEY/AGENT INFORMATION:
```

; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-073-064-2

Query Match   100.0%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCGGGGTGAACGGGGCCATT 27
       |||||||||||||||||||||||||||
Db 1 GGTGACTCACCGGGGTGAACGGGGCCATT 27

RESULT 2
US-10-683-516-3
; Sequence 3, Application US/10683516
; Publication No. US20040126364A1
; GENERAL INFORMATION:
; APPLICANT: Keeping, Hugh S
; APPLICANT: Reichner, Jonathan S
; TITLE OF INVENTION: Treatment for Bone Disorders
; FILE REFERENCE: 21486-028CON
; CURRENT APPLICATION NUMBER: US/10/683,516
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 09/507,239
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-683-516-3

Query Match   85.2%; Score 23; DB 17; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCGGGGTGAACGGGG 23
       ||||||||||||||||||||||||
Db 65 GGTGACTCACCGGGGTGAACGGGG 87

RESULT 3
US-10-171-311-193
; Sequence 193, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035

; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2623, 2624, 2625, 2626, 2627, 2628, 2629
; OTHER INFORMATION: n = A,T,C or G
US-10-171-311-193

Query Match   85.2%; Score 23; DB 14; Length 2657;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCGGGGTGAACGGGG 23
       ||||||||||||||||||||||||
Db 343 GGTGACTCACCGGGGTGAACGGGG 365

RESULT 4
US-10-029-386-19439/c
; Sequence 19439, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hansel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19439
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000392.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT: Q9RSK7, EVALUE 7.90e-02
; OTHER INFORMATION: NT HIT: AL603647.1, EVALUE 4.40e-01
; OTHER INFORMATION: EST_HUMAN HIT: AW069362.1, EVALUE 7.00e-38
US-10-029-386-19439

Query Match   67.4%; Score 18.2; DB 15; Length 132;
Best Local Similarity 87.0%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ACTCACCGGGGTGAACGGGGCCATT 27
       || |||||||||||||||||||
Db 55 ACCCTCGGGGTGAACGGGGCCATT 33

RESULT 5
US-10-027-632-26905
; Sequence 26905, Application US/10027632
; Publication No. US20020199371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 26905
LENGTH: 391
TYPE: DNA
ORGANISM: Human
US-10-027-632-26905

```
Qy   3 TGACTCACCGGGTGAACGGGGCA 25
       ||||||||||||| |||| |||||
Db  88 TGAGTCACCGGGAGAACGTGGCA 110
```

Query Match   67.4%; Score 18.2; DB 13; Length 391;
Best Local Similarity 87.0%; Pred. No. 81;
Matches   20; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

RESULT 6
US-10-027-632-26905
Sequence 26905, Application US10027632
Publication No. US20030240075A9
CURRENT FILING DATE: 2002-04-30
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
SEQ ID NO 26905
LENGTH: 391
TYPE: DNA
ORGANISM: Human
US-10-027-632-26905

Query Match   67.4%; Score 18.2; DB 15; Length 391;
Best Local Similarity 87.0%; Pred. No. 81;
Matches   20; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

RESULT 7
US-10-029-386-5683/c
Sequence 5683, Application US10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 5683
LENGTH: 545
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC000392.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.61
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.51
OTHER INFORMATION: EST HUMAN HIT: AW069362.1, EVALUE 8.00e-38
OTHER INFORMATION: SWISSPROT HIT: Q9V8M5, EVALUE 4.00e-02
US-10-029-386-5683

```
Qy   3 TGACTCACCGGGTGAACGGGGCA 25
       ||||||||||||| |||| |||||
Db  88 TGAGTCACCGGGAGAACGTGGCA 110
```

Query Match   67.4%; Score 18.2; DB 15; Length 545;
Best Local Similarity 87.0%; Pred. No. 81;
Matches   20; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

RESULT 8
US-10-165-800-23/c
Sequence 23, Application US10165800
Publication No. US20030092116A1
GENERAL INFORMATION:
APPLICANT: Chun, Miyoung
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: No. US2003092116A1e1 Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Adenylate Kinase, Phospholipid Scramblase-Like, DNA
TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
TITLE OF INVENTION: and ATPase-like Molecules and Uses Therefor
FILE REFERENCE: 35800/247838
CURRENT APPLICATION NUMBER: US/10/165,800
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 09/790,179
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,609
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,838
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/185,946
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,180
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,947
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/795,038
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/186,234
PRIOR FILING DATE: 2001-02-...

```
Qy   5 ACTCACCGGGTGAACGGGGCATT 27
       ||||| ||||||||||||||||||
Db 266 ACCCTCGGGTGAACGGGGCATT 244
```

Query Match   67.4%; Score 18.2; DB 15; Length 545;
Best Local Similarity 87.0%; Pred. No. 81;
Matches   20; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/781,677
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/181,705
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 23
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-165-800-23

Query Match   67.4%;   Score 18.2; DB 14; Length 774;
Best Local Similarity   87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           ||| |||||||||||||||||||
Db   596 ACCCTCGGGGTGAACGGGGCATT 574

RESULT 9
US-10-311-034-51/c
Sequence 51, Application US/10311034
Publication No. US20040023242A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: YUE, Henry
APPLICANT: LAL, Preeti
APPLICANT: BANDMAN, Olga
APPLICANT: BOROWSKY, Mark L.
APPLICANT: AU-YOUNG, Janice
APPLICANT: LU, Yan
APPLICANT: GANDHI, Ameena R.
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: CHAWLA, Narinder K.
APPLICANT: YAO, Monique G.
APPLICANT: LU, Dyung Aina M.
APPLICANT: GREENWALD, Sara R.
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: KEARNEY, Liam
APPLICANT: BURFORD, Neil
APPLICANT: NGUYEN, Danniel B.
APPLICANT: TANG, Y. Tom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: HE, Ann
APPLICANT: THORNTON, Michael
APPLICANT: HAFALIA, April
APPLICANT: ARVIZU, Chandra S.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: LO, Terence P.
APPLICANT: KHAN, Farrah A.
APPLICANT: RECIPON, Shirley A.
APPLICANT: AZIMZAI, Yalda
APPLICANT: POLICKY, Jennifer L.
APPLICANT: DING, Li
APPLICANT: GRETHER, Megan
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: BATRA, Sajeev
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: HUMAN KINASES
FILE REFERENCE: PI-0125 PCT
CURRENT APPLICATION NUMBER: US/10/311,034
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/212,073; 60/213,457; 60/215,651; 60/216,605; 60/218,372;
60/228,056
PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PERL Program

; SEQ ID NO 51
; LENGTH: 995
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 621299CB1

US-10-311-034-51

Query Match   67.4%;   Score 18.2; DB 16; Length 995;
Best Local Similarity   87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           ||| |||||||||||||||||||
Db   242 ACCCTCGGGGTGAACGGGGCATT 220

RESULT 10
US-09-764-868-257/c
Sequence 257, Application US/09764868
Patent No. US20020168711A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 257
LENGTH: 1223
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (78)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (307)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1163)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1191)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1193)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1203)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1218)
OTHER INFORMATION: n equals a,t,g, or c

US-09-764-868-257

Query Match   67.4%;   Score 18.2; DB 9; Length 1223;
Best Local Similarity   87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           ||| |||||||||||||||||||
Db   853 ACCCTCGGGGTGAACGGGGCATT 831

RESULT 11
US-10-141-634-1/c

Sequence 1, Application US/10141634
Publication No. US20030083365A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Xie, Qiongshu
APPLICANT: Abuin, Alejandro
APPLICANT: Walke, D. Wade
TITLE OF INVENTION: No. US200300083365A1e1 Human Kinases and Polynucleotides Encoding
FILE REFERENCE: LEX-0344-USA
CURRENT APPLICATION NUMBER: US/10/141,634
CURRENT FILING DATE: 2002-05-08
PRIOR APPLICATION NUMBER: US 60/289,727
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1440
TYPE: DNA
ORGANISM: homo sapiens

Query Match   67.4%;   Score 18.2;   DB 14;   Length 1440;
Best Local Similarity   87.0%;   Pred. No. 80;
Matches 20;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           |||||||||||||||||||||||
Db   812 ACCCTCGGGGTGAACGGGGGCATT 790

RESULT 12
US-10-165-800-21/c
Sequence 21, Application US/10165800
Publication No. US20030092116A1
GENERAL INFORMATION:
APPLICANT: Chun, Miyoung
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: No. US20030092116A1e1 Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Adenylate Kinase, Phospholipid Scramblase-Like, DNA
TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
TITLE OF INVENTION: and ATPase-Like Molecules and Uses Therefor
FILE REFERENCE: 35800/247838
CURRENT APPLICATION NUMBER: US/10/165,800
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 09/790,179
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,609
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 09/790,838
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/185,946
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 60/185,947
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,180
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 09/795,038
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/186,234
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/781,677
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/181,705
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 1452
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (219)...(995)

RESULT 13
US-10-094-749-1381/c
Sequence 1381, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1381
LENGTH: 1579
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-1381

Query Match   67.4%;   Score 18.2;   DB 14;   Length 1579;
Best Local Similarity   87.0%;   Pred. No. 80;
Matches 20;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           |||||||||||||||||||||||
Db   814 ACCCTCGGGGTGAACGGGGGCATT 792

RESULT 14
US-10-478-146-17/c
Sequence 17, Application US/10478146
Publication No. US20040203097A1
GENERAL INFORMATION:

Query Match   67.4%;   Score 18.2;   DB 15;   Length 1579;
Best Local Similarity   87.0%;   Pred. No. 80;
Matches 20;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy   5 ACTCACCGGGTGAACGGGGCATT 27
           |||||||||||||||||||||||
Db   841 ACCCTCGGGGTGAACGGGGGCATT 819

APPLICANT: INCYTE CORPORATION; YUE, Henry;
APPLICANT: LU, Dyung Aina M.; AZIMZAI, Yalda;
APPLICANT: DING, Li; LEE, Ernestine A.;
APPLICANT: HAFALIA, April J.A.; BECHA, Shanya D.;
APPLICANT: TANG, Y. Tom; LAL, Preeti G.;
APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
APPLICANT: RAMKUMAR, Jayalaxmi; ELLIOTT, Vicki S.;

```
US-10-104-047-1657
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   LENGTH: 1619
;   SEQ ID NO 1657
;   NUMBER OF SEQ ID NOS: 4096
;   SOFTWARE: PatentIn Ver. 2.1
;   PRIOR FILING DATE:
;   PRIOR APPLICATION NUMBER:
;   CURRENT FILING DATE: 2002-03-25
;   CURRENT APPLICATION NUMBER: US/10/104,047
;   FILE REFERENCE: H1-A0105
;   TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
;   APPLICANT: HELIX RESEARCH INSTITUTE
;   GENERAL INFORMATION:
;   Publication No. US2003023639A1
;   Sequence 1657, Application US/10104047
US-10-104-047-1657/c
RESULT 15

Qy          5 ACTCACCGGGTGAACGGGGGCATT 27
              |||||||||||||||||||||||||
Db        826 ACCCTCGGGGTGAACGGGGGCATT 804

Query Match        67.4%;  Score 18.2;  DB 15;  Length 1619;
Best Local Similarity   87.0%;  Pred. No. 80;
Matches   20;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;
```

```
US-10-478-146-17
;   OTHER INFORMATION: Incyte ID No: 72063274CB1
;   NAME/KEY: misc_feature
;   FEATURE:
;   ORGANISM: Homo sapiens
;   TYPE: DNA
;   LENGTH: 1581
;   SEQ ID NO 17
;   SOFTWARE: PERL Program
;   NUMBER OF SEQ ID NOS: 28
;   PRIOR FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/376,988
;   PRIOR FILING DATE: 2002-02-15
;   PRIOR APPLICATION NUMBER: US 60/357,675
;   PRIOR FILING DATE: 2001-12-19
;   PRIOR APPLICATION NUMBER: US 60/343,007
;   PRIOR FILING DATE: 2001-07-27
;   PRIOR APPLICATION NUMBER: US 60/308,183
;   PRIOR FILING DATE: 2001-07-19
;   PRIOR APPLICATION NUMBER: US 60/306,967
;   PRIOR FILING DATE: 2001-07-06
;   PRIOR APPLICATION NUMBER: US 60/303,418
;   PRIOR FILING DATE: 2001-06-15
;   PRIOR APPLICATION NUMBER: US 60/298,712
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 60/293,665
;   PRIOR FILING DATE: 2002-05-23
;   PRIOR APPLICATION NUMBER: PCT/US02/16634
;   PRIOR FILING DATE: 2003-11-18
;   CURRENT FILING DATE:
;   CURRENT APPLICATION NUMBER: US/10/478,146
;   FILE REFERENCE: PF-0995 USN
;   TITLE OF INVENTION: KINASES AND PHOSPHATASES
;   APPLICANT: CHIEN, David; JIN, Pei
;   APPLICANT: RICHARDSON, Thomas W.; MARQUIS, Joseph P.;
;   APPLICANT: BOROWSKY, Mark L.; ZEBARJADIAN, Yeganeh;
;   APPLICANT: KHAN, Farrah A.; BAUGHN, Mariah R.;
;   APPLICANT: GANDHI, Ameena R.; YAO, Monique G.;
;   APPLICANT: TRAN, Uyen K.; CHAWLA, Narinder K.;
;   APPLICANT: SWARNAKAR, Anita; DUGGAN, Brendan M.;
;   APPLICANT: ARVIZU, Chandra S.; LUO, Wen;
```

```
Qy          5 ACTCACCGGGTGAACGGGGGCATT 27
              ||||||||||||| |||||||||||
Db        826 ACCCTCGGGGTGAACGGGGGCATT 804

Query Match        67.4%;  Score 18.2;  DB 18;  Length 1581;
Best Local Similarity   87.0%;  Pred. No. 80;
Matches   20;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;
```

```
Search completed: November 29, 2004, 17:46:30
Job time : 4974 secs

Qy          5 ACTCACCGGGTGAACGGGGGCATT 27
              |||||||||||||||||||||||||
Db        883 ACCCTCGGGGTGAACGGGGGCATT 861
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.   GenCore version 5.1.6

Run on: November 29, 2004, 13:45:02 ; Search time 3814 Seconds
(without alignments)
287.067 Million cell updates/sec

Title: US-10-073-064-2
Perfect score: 27
Sequence: 1 GGTGACTCACCGGGTGAACGGGGGCATT 27

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 2027541876 5 residues

Total number of hits satisfying chosen parameters: 88617144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main.*

```
1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID		Description
1	27	100.0	27	1	PCT-US92-09353-2	Sequence 2, Appli
2	27	100.0	27	7	US-08-209-846B-2	Sequence 2, Appli
3	27	100.0	27	9	US-08-451-399A-2	Sequence 2, Appli
4	27	100.0	27	2	US-08-451-882-2	Sequence 2, Appli
5	27	100.0	27	16	US-09-181-419A-2	Sequence 2, Appli
6	27	100.0	27	46	US-10-073-064-2	Sequence 2, Appli
7	23	85.2	57	21	US-09-400-957-6	Sequence 6, Appli
8	23	85.2	57	21	US-09-400-957-6	Sequence 6, Appli
9	23	85.2	301	22	US-08-066-372-2	Sequence 2, Appli
10	23	85.2	422	22	US-09-528-409-59789	Sequence 59789,
11	23	85.2	422	39	US-09-933-524-59789	Sequence 59789, A
12	23	85.2	422	39	US-09-933-524A-59789	Sequence 59789, A
13	23	85.2	565	16	PCT-US02-22216-2	Sequence 2, Appli
14	23	85.2	565	16	PCT-US02-22216A-2	Sequence 2, Appli
15	23	85.2	597	22	US-09-507-239-3	Sequence 3, Appli
16	23	85.2	597	21	US-09-683-516-3	Sequence 3, Appli
17	23	85.2	1604	48	US-10-170-235-4174	Sequence 4174, Ap
18	23	85.2	1604	111	US-60-452-680-2065	Sequence 2065, Ap
19	23	85.2	1604	111	US-60-453-050-1290	Sequence 1290, Ap
20	23	85.2	1604	111	US-60-453-135-1290	Sequence 1290, Ap
21	23	85.2	1604	112	US-60-466-412-1290	Sequence 1290, Ap
22	23	85.2	1667	48	US-10-170-235-4178	Sequence 4178, Ap
23	23	85.2	1667	111	US-60-452-680-2059	Sequence 2059, Ap
24	23	85.2	1667	111	US-60-453-050-1284	Sequence 1284, Ap
25	23	85.2	1667	111	US-60-453-135-1284	Sequence 1284, Ap
26	23	85.2	1667	112	US-60-466-412-1284	Sequence 1284, Ap
27	23	85.2	1765	32	US-09-724-676-14621	Sequence 14621, A
28	23	85.2	1765	32	US-09-724-676A-14621	Sequence 14621, A
29	23	85.2	2571	32	US-09-724-676-14620	Sequence 14620, A
30	23	85.2	2571	32	US-09-724-676A-14620	Sequence 14620, A
31	23	85.2	2650	2	PCT-US02-08992-16	Sequence 16, Appl
32	23	85.2	2650	62	US-10-821-801-119	Sequence 119, App
33	23	85.2	2657	2	PCT-US02-18638A-193	Sequence 193, App
34	23	85.2	2657	48	US-10-171-311-193	Sequence 193, App
35	23	85.2	2657	54	US-10-473-974-100	Sequence 100, App
36	23	85.2	2677	30	US-09-699-997-11736	Sequence 11736, A
37	23	85.2	2677	30	US-09-699-998-10657	Sequence 10657, A
38	23	85.2	2677	31	US-09-716-472-4937	Sequence 4937, A
39	23	85.2	2677	31	US-09-716-972-4956	Sequence 4956, Ap
40	23	85.2	2696	32	US-09-724-676-14629	Sequence 14629, A
41	23	85.2	2696	32	US-09-724-676A-14629	Sequence 14629, A
42	23	85.2	2732	17	US-09-205-070-10104	Sequence 10104, A
43	23	85.2	2732	19	US-09-340-623-10104	Sequence 10104, A
44	23	85.2	2732	37	US-09-898-888-10104	Sequence 10104, A
45	23	85.2	2732	37	US-09-898-888A-10104	Sequence 10104, A

117:	/cgn2_6/ptodata/1/pna/US6051_COMB.seq:*
118:	/cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
119:	/cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
120:	/cgn2_6/ptodata/1/pna/US6053_COMB.seq:*
121:	/cgn2_6/ptodata/1/pna/US6054_COMB.seq:*
122:	/cgn2_6/ptodata/1/pna/US6055_COMB.seq:*
123:	/cgn2_6/ptodata/1/pna/US6056_COMB.seq:*
124:	/cgn2_6/ptodata/1/pna/US6057_COMB.seq:*
125:	/cgn2_6/ptodata/1/pna/US6058_COMB.seq:*
126:	/cgn2_6/ptodata/1/pna/US6059_COMB.seq:*

RESULT 1
PCT-US92-09353-2
; Sequence 2, Application PC/TUS9209353
; GENERAL INFORMATION:

ALIGNMENTS

APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Myogenic Vector Systems
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas D. Paul
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09353
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5379
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US92-09353-2

Query Match	100.0%;	Score 27;	DB 1;	Length 27;
Best Local Similarity	100.0%;			
Matches 27;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;

Qy   1 GGTGACTCACCGGGTGAACGGGGGCATT 27
          |||||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGTGAACGGGGGCATT 27

RESULT 2
US-08-209-846B-2
; Sequence 2, Application US/08209846B
; GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS AND METHODS
TITLE OF INVENTION: OF USE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,846B
FILING DATE: March 9, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/789,919
FILING DATE: November 6, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 204/302
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-209-846B-2

Query Match   100.0%; Score 27; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-209-846B-2

Qy 1 GGTGACTCACCGGGTGAACGGGGGCATT 27
        ||||||||||||||||||||||||||||
Db 1 GGTGACTCACCGGGTGAACGGGGGCATT 27

RESULT 3
US-08-451-399A-2
Sequence 2, Application US/08451399A
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS
TITLE OF INVENTION: AND METHOD OF USE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
        Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
        storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM MS-DOS
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,399A
FILING DATE: May 26, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,846
FILING DATE: March 9, 1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/789,919
FILING DATE: November 6, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Berkman, Charles S.
REGISTRATION NUMBER: 38,077
REFERENCE/DOCKET NUMBER: 212/136
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-451-399A-2

Query Match   100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-451-399A-2

Qy 1 GGTGACTCACCGGGTGAACGGGGGCATT 27
        ||||||||||||||||||||||||||||
Db 1 GGTGACTCACCGGGTGAACGGGGGCATT 27

RESULT 4
US-08-451-882-2
Sequence 2, Application US/08451882
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: DeMayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Expression Vector Systems and Methods of
TITLE OF INVENTION: Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth St.
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,882
FILING DATE: 26-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,846
FILING DATE: 09-MAR-1994
APPLICATION NUMBER: US 07/789,919
FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Knight, Matthew W.
REGISTRATION NUMBER: 36,846
REFERENCE/DOCKET NUMBER: 204/302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-489-1600
TELEFAX: 213-955-0440
TELEX: 67-3510
US-08-451-882-2

Query Match   100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-451-882-2

Qy 1 GGTGACTCACCGGGTGAACGGGGGCATT 27
        ||||||||||||||||||||||||||||
Db 1 GGTGACTCACCGGGTGAACGGGGGCATT 27

RESULT 5
US-09-181-419A-2
Sequence 2, Application US/0918141419A
GENERAL INFORMATION:
; APPLICANT: Valentis, Inc.
; APPLICANT: Schwartz, Robert J.
; APPLICANT: DeMayo, Franco J.
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS AND METHOD OF USE
; FILE REFERENCE: SB Meyer; Valentis 236/277 US
; CURRENT APPLICATION NUMBER: US/09/181,419A
; CURRENT FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 08/209,846
; PRIOR FILING DATE: 1994-03-09
; PRIOR APPLICATION NUMBER: 07/789,919
; PRIOR FILING DATE: 1991-11-06
; PRIOR APPLICATION NUMBER: 08/472,809
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-181-419A-2

Query Match   100.0%; Score 27; DB 16; Length 27;
Best Local Similarity   100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches   27; Conservative   0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCATT 27
         |||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGTGAACGGGGCATT 27

RESULT 6
US-10-073-064-2
Sequence 2, Application US/10073064
GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
;   DeMayo, Franco
;   O'Malley, Bert W.
; TITLE OF INVENTION: Expression Vector Systems and Methods of
;   Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth St.
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/073,064
; FILING DATE: 12-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,265
; FILING DATE: 09-May-1995
; APPLICATION NUMBER: 08/368,776
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/789,919
; FILING DATE: 06-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846

; REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-073-064-2

Query Match   100.0%; Score 27; DB 46; Length 27;
Best Local Similarity   100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches   27; Conservative   0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCATT 27
         |||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGTGAACGGGGCATT 27

RESULT 7
US-09-400-967-6
Sequence 6, Application US/09400967
GENERAL INFORMATION:
; APPLICANT: Adams, John S.
; TITLE OF INVENTION: Nucleic Acids Encoding Vitamin D Response Element
; TITLE OF INVENTION: Binding Proteins, Products Related Thereto, and Methods
; TITLE OF INVENTION: of Using Same
; FILE REFERENCE: P-CE 3489
; CURRENT APPLICATION NUMBER: US/09/400,967
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 6
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-400-967-6

Query Match   85.2%; Score 23; DB 21; Length 57;
Best Local Similarity   100.0%; Pred. No. 8.2; Mismatches 0; Indels 0; Gaps 0;
Matches   23; Conservative   0;

Qy   1 GGTGACTCACCGGGTGAACGGGG 23
         |||||||||||||||||||||||
Db   3 GGTGACTCACCGGGTGAACGGGG 25

RESULT 8
US-09-400-967-6/c
Sequence 6, Application US/09400967
GENERAL INFORMATION:
; APPLICANT: Adams, John S.
; TITLE OF INVENTION: Nucleic Acids Encoding Vitamin D Response Element
; TITLE OF INVENTION: Binding Proteins, Products Related Thereto, and Methods
; TITLE OF INVENTION: of Using Same
; FILE REFERENCE: P-CE 3489
; CURRENT APPLICATION NUMBER: US/09/400,967
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 6
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-400-967-6

Query Match   85.2%; Score 23; DB 21; Length 57;
Best Local Similarity   100.0%; Pred. No. 8.2;

```
Matches  23;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  GGTGACTCACCGGGTGAACGGGG  23
         |||||||||||||||||||||||
Db   56  GGTGACTCACCGGGTGAACGGGG  34
```

RESULT 9
US-08-066-372-2
; Sequence 2, Application US/08066372
; GENERAL INFORMATION:
; APPLICANT: Li, Yi-Ping
; APPLICANT: Stashenko, Philip
; TITLE OF INVENTION: TNF' RESPONSIVE ELEMENT, TNF'-INDUCED DNA-BINDING
; TITLE OF INVENTION: PROTEIN AND METHOD OF INHIBITING TNF' DOWN-REGULATION OF GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; ZIP: 02173
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/066,372
; FILING DATE: 19930520
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: FDC93-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

```
Query Match         85.2%;  Score 23;  DB 5;  Length 301;
Best Local Similarity   100.0%;  Pred. No. 8.8;
Matches  23;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

US-08-066-372-2

Qy    1  GGTGACTCACCGGGTGAACGGGG  23
         |||||||||||||||||||||||
Db  240  GGTGACTCACCGGGTGAACGGGG  262
```

RESULT 10
US-09-528-409-59789
; Sequence 59789, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; TITLE OF INVENTION: From Various Libraries
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 59789
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(422)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-59789

```
Qy    1  GGTGACTCACCGGGTGAACGGGG  23
         |||||||||||||||||||||||
Db  330  GGTGACTCACCGGGTGAACGGGG  352
```

```
Query Match         85.2%;  Score 23;  DB 22;  Length 422;
Best Local Similarity   100.0%;  Pred. No. 8.9;
Matches  23;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 11
US-09-933-524-59789
; Sequence 59789, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; TITLE OF INVENTION: From Various Libraries
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 59789
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(422)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524-59789

```
Query Match         85.2%;  Score 23;  DB 39;  Length 422;
Best Local Similarity   100.0%;  Pred. No. 8.9;
Matches  23;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  GGTGACTCACCGGGTGAACGGGG  23
         |||||||||||||||||||||||
Db  330  GGTGACTCACCGGGTGAACGGGG  352
```

RESULT 12
US-09-933-524A-59789
; Sequence 59789, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; TITLE OF INVENTION: From Various Libraries
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409

```
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 59789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(422)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-59789

Qy          LENGTH: 422

Query Match              85.2%;  Score 23;  DB 39;  Length 422;
Best Local Similarity   100.0%;  Pred. No. 8.9;
Matches   23;  Conservative    0;  Mismatches     0;  Indels    0;  Gaps    0;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db        330 GGTGACTCACCGGGTGAACGGGG 352

RESULT 13
PCT-US02-22216-2
; Sequence 2, Application PC/TUS0222216
; GENERAL INFORMATION:
; APPLICANT: The University of Virginia Patent Foundation
; APPLICANT: Chung, Leland
; APPLICANT: Yeung, Fan
; TITLE OF INVENTION: Super Osteocalcin Promoter for the Treatment of Calcified Tumors
; FILE REFERENCE: 00704-02
; CURRENT APPLICATION NUMBER: PCT/US02/22216
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,360
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-22216-2

Qy          LENGTH: 565

Query Match              85.2%;  Score 23;  DB 2;  Length 565;
Best Local Similarity   100.0%;  Pred. No. 9;
Matches   23;  Conservative    0;  Mismatches     0;  Indels    0;  Gaps    0;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db         74 GGTGACTCACCGGGTGAACGGGG 96

RESULT 14
PCT-US02-22216A-2
; Sequence 2, Application PC/TUS0222216A
; GENERAL INFORMATION:
; APPLICANT: The University of Virginia Patent Foundation
; APPLICANT: Chung, Leland W. K.
; APPLICANT: Yeung, Fan
; TITLE OF INVENTION: Super Osteocalcin Promoter for the Treatment of Calcified Tumors
; FILE REFERENCE: 00704-02
; CURRENT APPLICATION NUMBER: PCT/US02/22216A
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,360
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-22216A-2

Qy

Query Match              85.2%;  Score 23;  DB 2;  Length 565;
Best Local Similarity   100.0%;  Pred. No. 9;
Matches   23;  Conservative    0;  Mismatches     0;  Indels    0;  Gaps    0;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db         74 GGTGACTCACCGGGTGAACGGGG 96

RESULT 15
US-09-507-239-3
; Sequence 3, Application US/09507239
; GENERAL INFORMATION:
; APPLICANT: Keeping, Hugh S
; APPLICANT: Reichner, Jonathan S
; TITLE OF INVENTION: Treatment for Bone Disorders
; FILE REFERENCE: 21486-028
; CURRENT APPLICATION NUMBER: US/09/507,239
; CURRENT FILING DATE: 2001-02-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 597
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-507-239-3

Qy

Query Match              85.2%;  Score 23;  DB 22;  Length 597;
Best Local Similarity   100.0%;  Pred. No. 9;
Matches   23;  Conservative    0;  Mismatches     0;  Indels    0;  Gaps    0;

Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db         65 GGTGACTCACCGGGTGAACGGGG 87

Search completed: November 29, 2004, 16:15:10
Job time : 3816 secs
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: November 29, 2004, 13:45:01 ; Search time 2981 Seconds
(without alignments)
330.048 Million cell updates/sec

Title: US-10-073-064-2
Perfect score: 27
Sequence: 1 GGTGACTCACCGGGTGAACGGGGCATT 27
Scoring table: IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0
Searched: 32822875 seqs, 1821865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
    1: gb_est1:*
    2: gb_est2:*
    3: gb_htc:*
    4: gb_est3:*
    5: gb_est4:*
    6: gb_est5:*
    7: gb_est6:*
    8: gb_gss1:*
    9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	85.2	436	7	R34738	R34738 yg61d05.r1
2	19.6	72.6	495	1	AL904250	AL904250 AL904250
3	19.6	72.6	539	1	AL927557	AL927557 AL927557
4	19.6	72.6	776	9	CR131630	CR131630 Forward s
5	19.2	71.1	1150	2	BF026229	BF026229 601672804
6	19	70.4	313	6	CD587090	CD587090 RK039A3H1
7	19	70.4	352	6	CD587012	CD587012 RK039A1H1
8	19	70.4	529	1	AI765212	AI765212 wi72f04.x
9	19	70.4	699	8	BH474984	BH474984
10	19	70.4	699	8	BH976919	BH976919 cdh11f07.
11	19	70.4	725	8	BH517746	BH517746 BCGYE61F
12	19	70.4	842	6	CA496206	CA496206 AGENCOURT
13	18.6	68.9	716	7	CF879332	CF879332 trico80xd
14	18.6	68.9	771	5	CB905760	CB905760 trico80xd
15	18.6	68.9	790	5	BX368969	BX368969
16	18.6	68.9	855	9	CN500A99	CN500A99 Drosophil
17	18.6	68.9	866	7	CK865132	CK865132 AGENCOURT
18	18.6	68.9	976	4	BG333940	BG333940 602460282
19	18.6	68.9	1079	4	BM800594	BM800594 AGENCOURT
20	18.6	68.9	1677	3	CR668008	CR668008 Tetraodon
21	18.2	67.4	406	7	CK554185	CK554185 rswla0_01
22	18.2	67.4	450	8	BZ096975	BZ096975 CH230-142
23	18.2	67.4	463	2	AW069362	AW069362 cr44g08.x
24	18.2	67.4	540	4	BJ689916	BJ689916 BJ689916
25	18.2	67.4	557	9	CE248534	CE248534 tigr-gss-
26	18.2	67.4	622	7	CK535339	CK535339 rswgb0_00
27	18.2	67.4	675	9	CC711406	CC711406 OGWET58TH
28	18.2	67.4	802	9	CN502ADH	CN502ADH Tetraodon
29	18.2	67.4	852	2	BE545517	BE545517 601076509
30	18.2	67.4	2567	9	BC037402	BC037402 Homo sapi
31	18	66.7	221	9	CE046711	CE046711 tigr-gss-
32	18	66.7	261	1	AA249500	AA249500 j811.seq
33	18	66.7	321	1	AL374743	AL374743 AL374743
34	18	66.7	326	1	AL376677	AL376677 AL376677
35	18	66.7	335	9	CC645393	CC645393 AV659678
36	18	66.7	360	1	AV659678	AV659678 AV659678
37	18	66.7	418	1	AL376678	AL376678 MtBB25D06
38	18	66.7	419	1	AL388075	AL388075 MtBC46E02
39	18	66.7	427	1	AL388076	AL388076 MtBC46E02
40	18	66.7	436	4	BI271584	BI271584 NF100G12F
41	18	66.7	461	1	AL376435	AL376435 MtBB23H04
42	18	66.7	464	1	AL376436	AL376436 MtBB23H04
43	18	66.7	468	5	BX472073	BX472073 DKFZp686B
44	18	66.7	472	1	AL370632	AL370632 MtBA59A04
45	18	66.7	474	5	BQ147306	BQ147306 NF039C04F

ALIGNMENTS

RESULT 1
R34738

```
LOCUS       R34738                   436 bp    mRNA       linear   EST 02-MAY-1995
DEFINITION  yg61d05.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:37522 5', mRNA sequence.
ACCESSION   R34738
VERSION     R34738.1  GI:791639
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2249
            High quality sequence stops: 329  Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2249    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 329.
FEATURES             Location/Qualifiers
     source          1..436
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:410063"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:37522"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares infant brain 1NIB"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site 1: Not
                     I; Site 2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
```

ORIGIN
```
AACTGGAAGAATTCGGCGCCCGGAGGAATTTTTTTTTTTTTTTTTTT 3');
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

Qy   1 GGTGACTCACCGGGTGAACGGGGG 23
             ||||||||||||||||||||||||
Db   304 GGTGACTCACCGGGTGAACGGGGG 326

RESULT 2
Query Match   85.2%; Score 23; DB 7; Length 436;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS   AL904250/c
DEFINITION AL904250 PJR-21+22 Danio rerio cDNA clone 013-E07-1, mRNA sequence.
ACCESSION   AL904250
VERSION   AL904250.1 GI:23166809
KEYWORDS   EST.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 495)
  AUTHORS   Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
            Wang,W., Wen,Z. and Peng,J.
  TITLE   15000 unique zebrafish EST clusters and their future use in
            microarray for profiling gene expression patterns during
            embryogenesis
  JOURNAL   Genome Res. 13 (3), 455-466 (2003)
   MEDLINE 22505427
   PUBMED   12618376
COMMENT   Contact: Peng J
            Lab of Functional Genomics
            Institute of Molecular and Cell Biology
            30 Medical Drive, Singapore, 117609, Singapore
            Email: pengjr@imcb.a-star.edu.sg
            Clone requests: info@openbiosystems.com
            Open Biosystems
            6705 Odyssey Drive, Huntsville, AL 35806.
FEATURES   Location/Qualifiers
     source   1..495
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="local wildtype"
                     /db_xref="taxon:7955"
                     /clone="013-E07-1"
                     /tissue_type="whole embryo or fish"
                     /dev_stage="mixed stages"
                     /clone_lib="PJR-21+22"

ORIGIN

Qy   1 GGTGACTCACCGGGTGAACGGGGCAT 26
             ||||||||||||||||||||||||||
Db   261 GGTGACCACCACCGGGTAACCGGGGAAT 236

RESULT 3
Query Match   72.6%; Score 19.6; DB 1; Length 539;
Best Local Similarity 84.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

LOCUS   AL927557/c
DEFINITION AL927557 PJR-21+22 Danio rerio cDNA clone 191-E01-1, mRNA sequence.
ACCESSION   AL927557
VERSION   AL927557.1 GI:23198469

ORIGIN

Query Match

KEYWORDS   EST.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 539)
  AUTHORS   Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
            Wang,W., Wen,Z. and Peng,J.
  TITLE   15000 unique zebrafish EST clusters and their future use in
            microarray for profiling gene expression patterns during
            embryogenesis
  JOURNAL   Genome Res. 13 (3), 455-466 (2003)
   MEDLINE 22505427
   PUBMED   12618376
COMMENT   Contact: Peng J
            Lab of Functional Genomics
            Institute of Molecular and Cell Biology
            30 Medical Drive, Singapore, 117609, Singapore
            Email: pengjr@imcb.a-star.edu.sg
            Clone requests: info@openbiosystems.com
            Open Biosystems
            6705 Odyssey Drive, Huntsville, AL 35806.
FEATURES   Location/Qualifiers
     source   1..539
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="local wildtype"
                     /db_xref="taxon:7955"
                     /clone="191-E01-1"
                     /tissue_type="whole embryo or fish"
                     /dev_stage="mixed stages"
                     /clone_lib="PJR-21+22"

ORIGIN

Qy   1 GGTGACTCACCGGGTGAACGGGGCAT 26
             |||||||||||||||||| ||||| ||
Db   392 GGTGACCACCACCGGGTAACCGGGGAAT 367

RESULT 4
Query Match   72.6%; Score 19.6; DB 1; Length 776;
Best Local Similarity 84.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

LOCUS   CR131630/c
DEFINITION CR131630 Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP255023, genomic survey sequence.
ACCESSION   CR131630
VERSION   CR131630.1 GI:49870983
KEYWORDS   GSS; Genome survey sequence; MICER.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 776)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK, http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
     source   1..776
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP259023"
                     /clone_lib="MHPP"

ORIGIN

Query Match   72.6%; Score 19.6; DB 9; Length 776;
```

(Additional result-header fragments visible on left margin:)

RESULT 2
LOCUS       AL904250
DEFINITION  AL904250     495 bp    mRNA    linear    EST 02-JUL-2004

RESULT 3
LOCUS       AL927557     539 bp    mRNA    linear    EST 06-JUL-2004

RESULT 4
LOCUS       CR131630     776 bp    DNA     linear    GSS 06-JUL-2004

RESULT 5
BF026229

```
Qy    2 GTGACTCACCGGGGTGAACGGGCATT 27
        ||| |||||||| |||||||| ||||
Db  207 GTAACTCACCGTGTGAACCAGGCATT 182
```

Best Local Similarity 84.6%; Pred. No. 3.1e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db   BF026229

```
LOCUS       BF026229       1150 bp    mRNA    linear   EST 10-OCT-2000
DEFINITION  601672804F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3955633 5',
            mRNA sequence.
ACCESSION   BF026229
VERSION     BF026229.1  GI:10733941
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1150)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM830 row: h column: 02
            High quality sequence stop: 393.
FEATURES             Location/Qualifiers
     source          1..1150
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3955633"
                     /tissue_type="melanocic melanoma"
                     /clone_lib="NIH_MGC_20"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
```

FEATURES   Location/Qualifiers
     source   1..313
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="RK039A3H11"
                     /dev_stage="mature"
                     /clone_lib="Zebrafish Kidney Marrow cDNA library"
                     /note="Organ: kidney; Vector: pBS-CMV; Site_1: XhoI;
                     Site_2: EcoRI; Total RNA was extracted from the kidney
                     tissues of mature zebrafish. The poly (A)+ RNA fraction
                     was separated from total RNA by oligo (dT) cellulose
                     chromatography. Library was initially constructed in the
                     lambdaZAP Express vector (Stratagene) and in vivo excised
                     into pBS-CMV vector."
ORIGIN

RESULT 7
CD587012

```
Qy    1 GGTGACTCACCGGGGTGAACGGGGCATT 27
        | |||||  ||| |||||| |||||||||
Db  197 GTTGACCGCCCCGGGTGTGAACGGGCATT 223
```

Query Match   Best Local Similarity 70.4%; Score 19; DB 6; Length 313;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db   197 GTTGACCGCCCCGGGTGTGAACGGGCATT 223

```
LOCUS       CD587012       352 bp    mRNA    linear   EST 16-JUN-2003
DEFINITION  RK039A1H11.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA
            clone RK039A1H11 5', mRNA sequence.
ACCESSION   CD587012
VERSION     CD587012.1  GI:31768364
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 352)
AUTHORS     Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Zhang,G.W.,
            Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Look,A.T.,
            Zon,L.I. and Chen,Z.
TITLE       Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue
JOURNAL     Unpublished (2003)
COMMENT     Contact: Chen Z.
            State Key Lab for Medical Genomics
            Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
            Shanghai Second Medical University
            197 Rui Jin Road II, Shanghai 200025, P. R. China
            Tel: 86-21-64740490
            Fax: 86-21-64743206
            Email: zchen@stn.sh.cn
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..352
```

RESULT 6
CD587090

```
Qy    2 GTGACTCACCGGGGTGAACGGGGGCA 25
        ||||||||||||||| ||||||| ||
Db  562 GAGACTCCCGGGGTGAACCGGGCA 585
```

Query Match   Best Local Similarity 87.5%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db   CD587090

```
LOCUS       CD587090       313 bp    mRNA    linear   EST 16-JUN-2003
DEFINITION  RK039A3H11.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA
            clone RK039A3H11 5', mRNA sequence.
ACCESSION   CD587090
VERSION     CD587090.1  GI:31768442
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 313)
AUTHORS     Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Zhang,G.W.,
            Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Look,A.T.,
            Zon,L.I. and Chen,Z.
TITLE       Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue
JOURNAL     Unpublished (2003)
COMMENT     Contact: Chen Z.
            State Key Lab for Medical Genomics
            Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
            Shanghai Second Medical University
            197 Rui Jin Road II, Shanghai 200025, P. R. China
            Tel: 86-21-64740490
            Fax: 86-21-64743206
            Email: zchen@stn.sh.cn
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..313
```

RESULT 8

```
Qy      1 GGTGACTCACCGGGTGAACGGGGGCATT 27
          |||||| ||||||| ||||||||||||||
Db    197 GTTGACGCCCCGGTGTGAACGGGCGCATT 223
```

Query Match   70.4%;   Score 19;   DB 6;   Length 352;
Best Local Similarity   81.5%;   Pred. No. 5.5e+02;
Matches   22;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

```
ACCESSION   AI765212
VERSION     AI765212.1   GI:5231721
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 461.
FEATURES             Location/Qualifiers
     source          1..529
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2398879"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid12"
                     /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatina Bonaldo."
```

RESULT 8

LOCUS   AI765212   529 bp   mRNA   linear   EST 28-JUN-1999
DEFINITION   wi72f04.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398879 3'
   similar to SW:FACC_HUMAN Q00597 FANCONI ANEMIA GROUP C PROTEIN ;,
   mRNA sequence.

RESULT 8 (ORIGIN block)

ORIGIN

Query Match   70.4%;   Score 19;   DB 6;   Length 352;
Best Local Similarity   81.5%;   Pred. No. 5.5e+02;
Matches   22;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

```
     /organism="Danio rerio"
     /mol_type="mRNA"
     /db_xref="taxon:7955"
     /clone="RK039A1H11"
     /dev_stage="mature"
     /clone_lib="Zebrafish Kidney Marrow cDNA library"
     /note="Organ: Kidney; Vector: pBS-CMV; Site_1: XhoI;
     Site_2: EcoRI; Total RNA was extracted from the kidney
     tissues of mature zebrafish. The poly (A)+ RNA fraction
     was separated from total RNA by oligo (dT) cellulose
     chromatography. Library was initially constructed in the
     lambdaZAP Express vector (Stratagene) and in vivo excised
     into pBS-CMV vector."
```

RESULT 9

```
Qy      1 GGTGACTCACCGGGTGAACGGGGGCATT 27
          |||||||| ||||||||||||||| |||
Db    124 GGTGACTCTCCTGGTTATCTGGGCATT 98
```

Query Match   70.4%;   Score 19;   DB 1;   Length 529;
Best Local Similarity   81.5%;   Pred. No. 5.7e+02;
Matches   22;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

```
ACCESSION   BH474984
VERSION     BH474984.1   GI:17683095
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..699
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGFA32"
                     /clone_lib="BOGF"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
```

LOCUS   BH474984   699 bp   DNA   linear   GSS 13-DEC-2001
DEFINITION   BOGFA322F BOGF Brassica oleracea genomic clone BOGFA32, genomic
   survey sequence.

ORIGIN

RESULT 10

```
Qy      1 GGTGACTTCGCCGGGGAGAAAGGGGCTTT 27
          |||||| |||||| |||||||||| |||
Db    191 GGTGATTCGCCGCGGAGAAAGGGGCTTT 217
```

Query Match   70.4%;   Score 19;   DB 8;   Length 699;
Best Local Similarity   81.5%;   Pred. No. 5.7e+02;
Matches   22;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

```
ACCESSION   BH976919
VERSION     BH976919.1   GI:23459922
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
```

LOCUS   BH976919   699 bp   DNA   linear   GSS 02-OCT-2002
DEFINITION   odh11f07.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
   sequence.

TITLE   Nash,W., Rabinowicz,P.D. and Wilson,R.K.
          Whole genome shotgun reads from Brassica oleracea
JOURNAL   Unpublished (2002)
COMMENT   Contact: Richard K. Wilson
          Genome Sequencing Center
          Washington University School of Medicine
          Email: submissions@watson.wustl.edu
          Plate: odh11 row: f column: 07
          Seq primer: -28RPpOT reverse
          Class: shotgun
          High quality sequence start: 17
          High quality sequence stop: 551.
FEATURES   Location/Qualifiers
     source   1..699
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3712"
                     /clone_lib="B.oleracea002"
                     /note="Vector: pOTw13; Whole genome shotgun library from
                     flowering buds. DNA was purified from a crude nuclear
                     prep using Brassica oleracea TO100DH3 buds provided by
                     Thomas Osborn at the University of Wisconsin. Genomic
                     DNA was provided by Pablo Rabinowicz (CSHL) and the
                     shotgun library prepared at Washington University Genome
                     Sequencing Center."

ORIGIN
     Query Match   70.4%; Score 19; DB 8; Length 699;
     Best Local Similarity 81.5%; Pred. No. 5.7e+02;
     Matches   22; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   1 GGTGGACTCACCGGGTGAACGGGGCATT 27
          |||||| |||||| ||| |||||||||||
Db   426 GGTGGATTCGCTGGGAGAACGGGGCTTT 452

RESULT 11
LOCUS   BH517746
DEFINITION BOGYE61TF BOGY Brassica oleracea genomic survey sequence.
ACCESSION   BH517746
VERSION   BH517746.1 GI:17725836
KEYWORDS   GSS.
SOURCE   Brassica oleracea
  ORGANISM Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 725)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE   Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT   Other GSSs: BOGYE61TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            /clone="BOGYE61"
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..725
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO100DH3"
                     /db_xref="taxon:3712"
                     /clone_lib="BOGY"
                     /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"

LOCUS   BH517746   725 bp   DNA   linear   GSS 13-DEC-2001

ORIGIN
     Query Match   70.4%; Score 19; DB 8; Length 725;
     Best Local Similarity 81.5%; Pred. No. 5.7e+02;
     Matches   22; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   1 GGTGGACTCACCGGGTGAACGGGGCATT 27
          |||||| |||||| ||| |||||||||||
Db   226 GGTGGATTCGCTGGGAGAACGGGGCTTT 200

RESULT 12
LOCUS   CA496206
DEFINITION AGENCOURT_10811996 NCI_CGAP_ZKidl Danio rerio cDNA clone
            IMAGE:6791831 5', mRNA sequence.
ACCESSION   CA496206
VERSION   CA496206.1 GI:24959287
KEYWORDS   EST.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 842)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14296 row: b column: 22
            High quality sequence stop: 641.
FEATURES   Location/Qualifiers
     source   1..842
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6791831"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZKidl"
                     /note="Organ: kidney; Vector: pCMV-SPORT6.1; Primer:
                     EcoRV; Site 2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.8 kb. Constructed by J.
                     Wang (Research Genetics, Invitrogen Corp) from tissue
                     donated by L. Zon (Harvard University). Note: this is a
                     NCI_CGAP Library."

LOCUS   CA496206   842 bp   mRNA   linear   EST 14-NOV-2002

ORIGIN
     Query Match   70.4%; Score 19; DB 8; Length 842;
     Best Local Similarity 81.5%; Pred. No. 5.8e+02;
     Matches   22; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   1 GGTGGACTCACCGGGTGAACGGGGCATT 27
          |||||| |||||| ||| |||||||||||
Db   110 GGTGGATTCGCGCCCCGTGTGAATGGGCATT 84

RESULT 13
LOCUS   CF879332
DEFINITION tric080xd08.b1 T.reesei mycelial culture, Version 6 October 2003
            Hypocrea jecorina cDNA clone tric080xd08, mRNA sequence.
ACCESSION   CF879332
VERSION   CF879332.1 GI:38134014
KEYWORDS   EST.
SOURCE   Hypocrea jecorina (anamorph: Trichoderma reesei)

LOCUS   CF879332   716 bp   mRNA   linear   EST 31-OCT-2003

```
ORGANISM  Hypocrea jecorina
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
          Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE  1 (bases 1 to 716)
AUTHORS   Diener,S.E., Dunn-Coleman,N., Foreman,P., Houfek,T.D.,
          Teunissen,P.J.M., van Solingen,P., Dankmeyer,L., Mitchell,T.K.,
          Ward,M. and Dean,R.A.
TITLE     Characterization of the protein processing and secretion pathways
          in a comprehensive set of expressed sequence tags from Trichoderma
          reesei
JOURNAL   FEMS Microbiol. Lett. 230 (2), 275-282 (2004)
COMMENT   Contact: Ralph A. Dean
          Fungal Genomics Laboratory
          North Carolina State University
          Campus Box 7251, Raleigh, NC 27695, USA
          Tel: 919-513-0020
          Fax: 919-513-0024
          Email: ralph_dean@ncsu.edu
          Seq primer: LT-F1 primer.
FEATURES             Location/Qualifiers
     source          1..716
                     /organism="Hypocrea jecorina"
                     /mol_type="mRNA"
                     /strain="QM6a"
                     /db_xref="taxon:51453"
                     /clone="tric080xd08"
                     /dev_stage="mycelia"
                     /clone_lib="T.reesei mycelial culture, Version 6 October
                     2003"
                     /note="Vector: pREP3Y; Site 1; Not I/Sal I; Mycelial
                     culture grown from 24 hrs to 6 days with varying Carbon
                     and Nitrogen sources and concentrations."
ORIGIN

Query Match             68.9%;  Score 18.6;  DB 7;  Length 716;
Best Local Similarity   84.0%;  Pred. No. 8.8e+02;
Matches  21;  Conservative   0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCA 25
        ||||| |||| ||||| |||| ||||
Db  43 GGTGACGCACGGGGTGCACGGGGAA 19


RESULT 14
CB905760/c
LOCUS       CB905760       771 bp    mRNA    linear   EST 02-JUL-2003
DEFINITION  tric080xd08 T.reesei mycelial culture, Version 3 april Hypocrea
            jecorina cDNA clone tric080xd08, mRNA sequence.
ACCESSION   CB905760
VERSION     CB905760.1  GI:30121418
KEYWORDS    EST.
SOURCE      Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM    Hypocrea jecorina
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE   1 (bases 1 to 771)
AUTHORS     Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
            Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
            Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
            Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.
TITLE       Transcriptional regulation of biomass-degrading enzymes in the
            filamentous fungus Trichoderma reesei
JOURNAL     J. Biol. Chem. 278 (34), 31988-31997 (2003)
MEDLINE     22803314
PUBMED      12788920
COMMENT     Contact: Pamela K. Foreman
            Genencor Intl.
            925 Page Mill Road, Palo Alto, CA 94304, USA
            Tel: (650) 846-7635
            Fax: (650) 621-7817
            Email: Pforeman@genencor.com
            Seq primer: LT-F1 primer.
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Hypocrea jecorina"
                     /mol_type="mRNA"
                     /strain="QM6a"
                     /clone="tric080xd08"
                     /db_xref="taxon:51453"
                     /dev_stage="mycelia"
                     /clone_lib="T.reesei mycelial culture, Version 3 april"
                     /note="Vector: pREP3Y; Site 1; Not I/Sal I; Mycelial
                     culture grown from 24 hrs to 6 days with varying Carbon
                     and Nitrogen sources and concentrations."
ORIGIN

Query Match             68.9%;  Score 18.6;  DB 6;  Length 771;
Best Local Similarity   84.0%;  Pred. No. 8.8e+02;
Matches  21;  Conservative   0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCA 25
        ||||| |||| ||||| |||| ||||
Db  98 GGTGACGCACGGGGTGCACGGGGAA 74


RESULT 15
BX368969/c
LOCUS       BX368969       790 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
            Homo sapiens cDNA clone CS0DL010YC17 5-PRIME, mRNA sequence.
ACCESSION   BX368969
VERSION     BX368969.2  GI:46573873
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 790)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30447801.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr;  Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer, Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            10183.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAE0292F02_AE02703_2&c=10183.
            r.
FEATURES             Location/Qualifiers
     source          1..790
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DL010YC17"
                     /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                     25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer, Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match             68.9%;  Score 18.6;  DB 5;  Length 790;
Best Local Similarity   84.0%;  Pred. No. 8.8e+02;
Matches  21;  Conservative   0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCA 25
```

Db   410 GGTGAATTCACCAGGGGGGCACGGGGCA 386
            |||||| ||||| ||||| ||| | ||||||||||

Search completed: November 29, 2004, 15:11:27
Job time : 2986 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: November 29, 2004, 13:45:01 ; Search time 1663 Seconds

Title: US-10-073-064-2
Perfect score: 27
Sequence: 1 GGTGACTCACCGGGTGAACGGGGCATT 27

Scoring table: IDENTITY_NUC
           Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues
           (without alignments)
           767.782 Million cell updates/sec

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
           Maximum Match 100%
           Listing first 45 summaries

Database :   GenEmbl:*
          1: gb_ba:*
          2: gb_btg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	ID	Description
1	27	100.0	27	6	AR009136	AR009136	Sequence
2	27	100.0	27	6	AR202296	AR202296	Sequence
3	23	85.2	2650	12	AB004306	AB004306	Artificia
c 4	23	85.2	57	6	AX100465	AX100465	Sequence
5	23	85.2	57	6	AX100465	AX100465	Sequence
6	23	85.2	597	6	E13404	E13404	Regulatory
6	23	85.2	597	6	E13404	E13404	Regulatory
7	23	85.2	831	9	AX226068	AX226068	Sequence
8	23	85.2	884	9	AY147065	AY147065	Homo sapi
9	23	85.2	1384	12	RATOSTEO	M34013	Homo sapien
10	23	85.2	1386	12	AB105218	AB105218	Mus muscu
c 11	23	85.2	1384	12	AB105219	AB105219	Mus muscu
12	23	85.2	2657	6	AX577978	AX577978	Sequence
13	23	85.2	164168	9	AL135927	AL135927	Human DNA
14	23	85.2	164179	9	AC007227	AC007227	Homo sapi
c 15	23	74.8	95687	5	BX511245	BX511245	Zebrafish
c 16	20.2	74.8	180506	2	BX005403	BX005403	Danio rer
c 17	20.2	74.8	275351	2	AC117883	AC117883	Danio rer
c 18	19.8	73.3	267431	2	AC142799	AC142799	Rattus no
c 19	19.6	72.6	13203	6	CO614669	CO614669	Sequence
c 20	19.6	72.6	154730	3	AC008184	AC008184	Drosophil
c 21	19.6	72.6	168062	2	AC012857	AC012857	Drosophil
c 22	19.6	72.6	172479	3	AC006590	AC006590	Drosophil
c 23	19.6	72.6	228644	10	AC102570	AC102570	Mus muscu
c 24	19.6	72.6	260027	3	AB003659	AB003659	Drosophil
c 25	19.6	72.6	263051	2	CR450811	CR450811	Danio rer
c 26	19.6	70.4	6846	6	CQ597323	CQ597323	Sequence
c 27	19	70.4	12631	6	CQ596429	CQ596429	Sequence
c 28	19	70.4	104334	5	AL606751	AL606751	Zebrafish
c 29	19	70.4	115940	5	AL119429	AL119429	Homo sapi
c 30	19	70.4	141855	7	BX322537	BX322537	Danio rer
c 31	19	70.4	152485	5	AL935211	AL935211	Zebrafish
c 32	19	70.4	153829	5	BX004864	BX004864	Zebrafish
c 33	19	70.4	154123	2	AC019562	AC019562	Drosophil
c 34	19	70.4	154348	5	BX950208	BX950208	Danio rer
c 35	19	70.4	158143	9	AL157384	AL157384	Human DNA
c 36	19	70.4	163815	9	AC138024	AC138024	Homo sapi
c 37	19	70.4	169986	2	BX957325	BX957325	Danio rer
c 38	19	70.4	169710	2	BX927124	BX927124	Danio rer
c 39	19	70.4	172741	5	AL954866	AL954866	Zebrafish
c 40	19	70.4	173298	3	AC064489	AC064489	Drosophil
c 41	19	70.4	176521	3	AC097725	AC097725	Drosophil
c 42	19	70.4	177426	3	CR456622	CR456622	Danio rer
c 43	19	70.4	181079	2	AC016667	AC016667	Homo sapi
c 44	19	70.4	184499	5	BX005201	BX005201	Zebrafish
c 45	19	70.4	184849	2	AC021209	AC021209	Homo sapi

ALIGNMENTS

RESULT 1
AR009136

LOCUS   AR009136   27 bp   DNA   linear
DEFINITION   Sequence 2 from patent US 5756264.
ACCESSION   AR009136
VERSION   AR009136.1   GI:3967941
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 27)
  AUTHORS   Schwartz,R.J., DeMayo,P.J. and O'Malley,B.W.
  TITLE   Expression vector systems and method of use
  JOURNAL   Patent: US 5756264-A 2 26-MAY-1998;
FEATURES   Location/Qualifiers
      source   1..27
                   /organism="unknown"
                   /mol_type="unassigned DNA"

Query Match   100.0%;   Score 27;   DB 6;   Length 27;
Best Local Similarity   100.0%;   Pred. No. 1.4;
Matches   27;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 GGTGACTCACCGGGTGAACGGGGCATT 27
         |||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGTGAACGGGGCATT 27

RESULT 2
AR202296

LOCUS   AR202296   27 bp   DNA   linear   PAT 20-APR-2002
DEFINITION   Sequence 2 from patent US 6361984.
ACCESSION   AR202296
VERSION   AR202296.1   GI:20256835
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 27)
  AUTHORS   Ciossek,T., Ullrich,A. and Millauer,B.
  TITLE   MDK1 polypeptides

```
JOURNAL     Patent: US 6361984-A 2 26-MAR-2002;
FEATURES
     source          1..27
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Qy   Query Match     100.0%;  Score 27;  DB 6;  Length 27;
     Best Local Similarity 100.0%;  Pred. No. 1.4;
     Matches 27;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGGGCATT 27
       |||||||||||||||||||||||||||||
Db   1 GGTGACTCACCGGGGTGAACGGGGGCATT 27


RESULT 3
AX100465

LOCUS       AX100465                 57 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 6 from Patent WO0121649.
ACCESSION   AX100465
VERSION     AX100465.1  GI:13619490
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Adams,J.S.
  TITLE     Nucleic acids encoding vitamin d response element binding proteins,
            products related thereto, and methods of using same
  JOURNAL   Patent: WO 0121649-A 6 29-MAR-2001;
            CEDARS-SINAI MEDICAL CENTER (US)
FEATURES             Location/Qualifiers
     source          1..57
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Qy   Query Match     85.2%;  Score 23;  DB 6;  Length 57;
     Best Local Similarity 100.0%;  Pred. No. 62;
     Matches 23;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGGG 23
       |||||||||||||||||||||||||
Db   3 GGTGACTCACCGGGGTGAACGGGGG 25


RESULT 4
AX100465/c

LOCUS       AX100465                 57 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 6 from Patent WO0121649.
ACCESSION   AX100465
VERSION     AX100465.1  GI:13619490
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Adams,J.S.
  TITLE     Nucleic acids encoding vitamin d response element binding proteins,
            products related thereto, and methods of using same
  JOURNAL   Patent: WO 0121649-A 6 29-MAR-2001;
            CEDARS-SINAI MEDICAL CENTER (US)
FEATURES             Location/Qualifiers
     source          1..57
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Qy   Query Match     85.2%;  Score 23;  DB 6;  Length 57;
     Best Local Similarity 100.0%;  Pred. No. 62;
     Matches 23;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGGG 23
       |||||||||||||||||||||||||
Db  56 GGTGACTCACCGGGGTGAACGGGGG 34


RESULT 5
E13404

LOCUS       E13404                  597 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  Regulatory region of human osteocalcin gene.
ACCESSION   E13404
VERSION     E13404.1  GI:3252209
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Negishi,Y., Kajiyama,N. and Kawaguchi,R.
  TITLE     RECOMBINANT VECTOR, CELL TRANSFORMED THEREBY, AND MEASUREMENT OF
            BONE METABOLISM CONTROL ACTIVITY OF VITAMIN D DERIVATIVE USING THE
  JOURNAL   Patent: JP 1997182586-A 1 15-JUL-1997;
COMMENT     OS   Homo sapiens (human)
            PN   JP 1997182586-A/1
            PD   15-JUL-1997
            PF   29-DEC-1995 JP 1995352817
            PI   NEGISHI YOICHI, KAJIYAMA NAOKI, KAWAGUCHI RYUJI PC
            C12N15/09,C12N5/10,C12P21/02,G01N33/82,(C12P21/02,C12R1:91); CC
            S R L;KK
            strandedness: Double;
            topology: linear;
            hypothetical: No;
            anti-sense: No;
            Key
FH
     source          1..597
                     /organism="Homo sapiens"
                     /cell_type="osteoblast"
                     /clone="OST-BETA-gal"
     promoter        1..597
                     /note="human osteocalcin gene promoter".
FEATURES             Location/Qualifiers
     source          1..597
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Qy   Query Match     85.2%;  Score 23;  DB 6;  Length 597;
     Best Local Similarity 100.0%;  Pred. No. 42;
     Matches 23;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 GGTGACTCACCGGGGTGAACGGGGG 23
       |||||||||||||||||||||||||
Db  65 GGTGACTCACCGGGGTGAACGGGGG 87


RESULT 6
AX226068

LOCUS       AX226068                597 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160083.
ACCESSION   AX226068
VERSION     AX226068.1  GI:15555393
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Keeping,H.S. and Reichner,J.S.
```

RESULT 7

```
Qy          1 GGTGACTCACCGGGGTGAACGGGGG 23
              |||||||||||||||||||||||||
Db         65 GGTGACTCACCGGGGTGAACGGGGG 87
```

ORIGIN

```
Db                                 /db_xref="taxon:9606"
                                   /mol_type="unassigned DNA"
              /organism="Homo sapiens"
      source  1..597
FEATURES             Location/Qualifiers
JOURNAL     Rhode Island Hospital, A Lifespan Partner (US)
            Patent: WO 0160983-A 3 23-AUG-2001;
TITLE       Treatment for bone disorders
```

```
Matches    23; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
Best Local Similarity  100.0%;  Pred. No. 42;
Query Match    85.2%;   Score 23;  DB 6;  Length 597;
```

ORIGIN

REFERENCE
TITLE   Direct Submission
AUTHORS   Li,L.
            1 (bases 1 to 831)

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM    Homo sapiens
SOURCE      Homo sapiens (human)
KEYWORDS    .
VERSION     AY147065.1  GI:22947853
ACCESSION   AY147065
DEFINITION  Homo sapiens osteocalcin gene, promoter sequence.
LOCUS       AY147065         831 bp    DNA     linear   PRI 16-SEP-2002
```

RESULT 8

```
Qy          1 GGTGACTCACCGGGGTGAACGGGGG 23
              |||||||||||||||||||||||||
Db        329 GGTGACTCACCGGGGTGAACGGGGG 351
```

ORIGIN

```
        gene  <1..>831
              /gene="osteocalcin"
    promoter  <1..>831
              /gene="osteocalcin"
```

```
FEATURES             Location/Qualifiers
      source  1..831
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
JOURNAL     Submitted (03-SEP-2002) Institute of Orthopaedics, XiJing Hospital,
            17# Changle West Road, Xi'an, Shaanxi 710032, P.R. China
```

```
Matches    23; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
Best Local Similarity  100.0%;  Pred. No. 39;
Query Match    85.2%;   Score 23;  DB 9;  Length 831;
```

```
LOCUS       RATOSTEO         884 bp    DNA     linear   PRI 28-FEB-2003
DEFINITION  Homo sapiens osteocalcin gene, promoter region and partial cds.
ACCESSION   M34013
VERSION     M34013.1  GI:205865
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 884)
AUTHORS     Morrison,N.A., Shine,J., Fragonas,J.C., Verkest,V., McMenemy,M.L.
            and Eisman,J.A.
TITLE       1,25-dihydroxyvitamin D-responsive element and glucocorticoid
            repression in the osteocalcin gene
JOURNAL     Science 246 (4934), 1158-1161 (1989)
MEDLINE     90069588
```

PUBMED   2588000

```
REFERENCE   2  (bases 1 to 884)
AUTHORS     Morrison,N.A., Shine,J., Fragonas,J.C., Verkest,V., McMenemy,M.L.
            and Eisman,J.A.
TITLE       Direct Submission
JOURNAL     Submitted (27-APR-1993) Garvan Institute of Medical Research, St.
            Vincents Hospital, Sydney, Australia.
```

```
FEATURES             Location/Qualifiers
      source  1..884
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
 misc_feature  1..881
              /note="contains promoter and 5'UTR"
        mRNA  <882..>884
              /product="osteocalcin"
         CDS  882..>884
              /codon_start=1
              /product="osteocalcin"
              /protein_id="AAF05868.1"
              /db_xref="GI:6223673"
              /translation="M"
```

ORIGIN

RESULT 9

```
Qy          1 GGTGACTCACCGGGGTGAACGGGGG 23
              |||||||||||||||||||||||||
Db        324 GGTGACTCACCGGGGTGAACGGGGG 346
```

```
Matches    23; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
Best Local Similarity  100.0%;  Pred. No. 39;
Query Match    85.2%;   Score 23;  DB 9;  Length 884;
```

```
LOCUS       AB105218        1384 bp    DNA     linear   SYN 24-JAN-2004
DEFINITION  Mus musculus transgenic DNA, hOSF1 transgene left flanking region,
            integrated synthetic HPoc-Hosf-1 transgene.
ACCESSION   AB105218
VERSION     AB105218.1  GI:34740152
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Tsujimura,A. and Hashimoto-Gotoh,T.
TITLE       Transgene integration sites of hOSF1 expression mouse
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 1384)
AUTHORS     Tsujimura,A. and Hashimoto-Gotoh,T.
TITLE       Direct Submission
JOURNAL     Submitted (10-MAR-2003) Atsushi Tsujimura, Kyoto Pref.Univ.of Med.,
            Dept.Biochem. & Mol.Genet. Res.Inst.for Neuro.Dis. &Geriat.;
            Kawaramachi-Hirokoji, Kajii-Cho 465, Kamigyo-Ku, Kyoto, Kyoto
            602-8566, Japan (E-mail:atsujii@koto.kpu-m.ac.jp,
            Tel:81-75-251-5850, Fax:81-75-251-5799)
```

```
FEATURES             Location/Qualifiers
      source  1..1384
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
      source  501..1384
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /transgenic
              /db_xref="taxon:32630"
 repeat_region  <1..500
              /standard_name="gamma satellite"
              /note="hOSF1 transgene left flanking region"
              /evidence=experimental
 misc_feature  501..>1384
              /note="HPoc-Hosf-1 transgene"
```

RESULT 10

```
Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db        840 GGTGACTCACCGGGTGAACGGGG 862
```

Best Local Similarity   85.2%;   Score 23;   DB 12;   Length 1384;
Matches 23;   Conservative   0;   Pred. No. 36;   Mismatches   0;   Indels   0;   Gaps   0;

Query Match

ORIGIN   /evidence=experimental

LOCUS   AB105219/c   1386 bp   DNA   linear   SYN 24-JAN-2004
DEFINITION Mus musculus transgenic DNA, integrated synthetic HPoc-Hosf-1
            transgene, hOSF1 transgene right flanking region.
ACCESSION   AB105219
VERSION   AB105219.1 GI:34740153
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1386)
  AUTHORS   Tsujimura,A. and Hashimoto-Gotoh,T.
  TITLE   Transgene integration sites of hOSF1 expression mouse
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1386)
  AUTHORS   Tsujimura,A. and Hashimoto-Gotoh,T.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-MAR-2003) Atsushi Tsujimura, Kyoto Pref.Univ. of Med.,
            Dept.Biochem. & Mol.Genet. Res.Inst.for Neuro.Dis.&Geriat.;
            Kawaramachi-Hirokoji, Kajii-Cho 465, Kamigyo-Ku, Kyoto, Kyoto
            602-8566, Japan (E-mail:atsuji@koto.kpu-m.ac.jp,
            Tel:81-75-251-5850, Fax:81-75-251-5799)
FEATURES   Location/Qualifiers
     source   1..1386
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /transgenic
     source   1..886
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="Hpoc-Hosf-1 transgene"
                     complement(<1..886)
     misc_feature   complement(887..>1386)
                     /note="HPoc-Hosf-1 transgene"
     repeat_region   complement(887..>1386)
                     /standard_name="gamma satellite"
                     /note="hOSF1 transgene right flanking region"
                     /evidence=experimental
```

RESULT 11

```
Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db        545 GGTGACTCACCGGGTGAACGGGG 523
```

Best Local Similarity   85.2%;   Score 23;   DB 12;   Length 1386;
Matches 23;   Conservative   0;   Pred. No. 36;   Mismatches   0;   Indels   0;   Gaps   0;

Query Match

ORIGIN

LOCUS   AB004306   2650 bp   DNA   linear   SYN 21-MAY-2003
DEFINITION Artificial construct HPoc-Hosf-1 transgene for osteoblast
            stimulating factor-1, complete cds.
ACCESSION   AB004306
VERSION   AB004306.2 GI:18642488
KEYWORDS   osteoblast stimulating factor-1.
SOURCE   synthetic construct
  ORGANISM synthetic construct

REFERENCE   1
  AUTHORS   Masuda,H., Tsujimura,A., Yoshioka,M., Arai,Y., Kuboki,Y., Mukai,T.,
            Nakamura,T., Tsuji,H., Nakagawa,M. and Hashimoto-Gotoh,T.
  TITLE   Bone mass loss due to estrogen deficiency is compensated in
            transgenic mice overexpressing human osteoblast stimulating
            factor-1
  JOURNAL   Biochem. Biophys. Res. Commun. 238 (2), 528-533 (1997)
   MEDLINE   97446018
   PUBMED   9299545
REFERENCE   2 (bases 1 to 2650)
  AUTHORS   Hashimoto-Gotoh,T.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-MAY-1997) Tamotsu Hashimoto-Gotoh, Kyoto Pref. Univ.
            of Medicine., RINDG, Dept. of Biochem. and Mol. Genet.;
            Kawaramachi-Hirokoji, Kyoto 602, Japan
            (E-mail:thg@koto.kpu-m.ac.jp, Tel:075-251-5850, Fax:075-251-5799)
COMMENT   On Feb 8, 2002 this sequence version replaced gi:2564039.
FEATURES   Location/Qualifiers
     source   1..2650
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="human osteocalcin promoter-driven human osf-1 gene
                     flanked by rabbit beta-globin 5'- and 3'- untranslated
                     regions; HPoc-Hosf-1 transgene used to generate transgenic
                     mice"
     promoter   15..882
                     /note="HPoc-Hosf-1 transgene
                     derived from human osteocalcin promoter"
     TATA_signal   823..828
     CDS   1542..2048
                     /note="HPoc-Hosf-1 transgene
                     derived from human osf-1 gene"
                     /codon_start=1
                     /transl_table=11
                     /product="osteoblast stimulating factor-1"
                     /protein_id="BAA22944.1"
                     /db_xref="GI:2564040"
                     /translation="MQAQQYQQQRRKFAAAFLAFIFILLAAVDTAEAGKKKERPEKKVKK
                     SDCGEWQWSVCVPTSGDCGLGTREGTRTGAECKQTMKTQRCKIPCNWKKQFGAECKYQ
                     FQAWGECDLNTALKTRTGSLKRALHNAECQKTVTISKPCGKLTKPKPQAESKKKKEG
                     KKQSKMLD"
```

RESULT 12

```
Qy          1 GGTGACTCACCGGGTGAACGGGG 23
              |||||||||||||||||||||||
Db        343 GGTGACTCACCGGGTGAACGGGG 365
```

Best Local Similarity   85.2%;   Score 23;   DB 12;   Length 2650;
Matches 23;   Conservative   0;   Pred. No. 32;   Mismatches   0;   Indels   0;   Gaps   0;

Query Match

ORIGIN

LOCUS   AX577978   2657 bp   DNA   linear   PAT 08-JAN-2003
DEFINITION Sequence 100 from Patent WO02081745.
ACCESSION   AX577978
VERSION   AX577978.1 GI:27647186
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J.,
            Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.
  TITLE   Genes involved in osteogenesis, and methods of use
  JOURNAL   Patent: WO 02081745-A1 100 17-OCT-2002;
            Aventis Pharma S.A. (FR)
FEATURES   Location/Qualifiers
     source   1..2657

ORIGIN

 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Qy Query Match 85.2%; Score 23; DB 6; Length 2657;
  Best Local Similarity 100.0%; Pred. No. 32;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCGGGGTGAACGGGG 23
    |||||||||||||||||||||||
Db 343 GGTGACTCACCGGGGTGAACGGGG 365

RESULT 13
LOCUS   AL135927 164168 bp DNA linear PRI 30-SEP-2000
DEFINITION Human DNA sequence from clone RP11-54H19 on chromosome 1, complete
    sequence.
ACCESSION AL135927
VERSION AL135927.14 GI:10443354
KEYWORDS HTG.
SOURCE   Homo sapiens (human)
 ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164168)
 AUTHORS Howden,P.
 TITLE   Direct Submission
 JOURNAL Submitted (27-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
    CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
    requests: clonerequest@sanger.ac.uk
COMMENT   On Oct 1, 2000 this sequence version replaced gi:10185404.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission
    corresponding to the overlapping clone, as we submit sequences with
    only a small overlap as described above.
    This sequence has been finished according to sequence map criteria
    as follows. An attempt is made to resolve all sequencing problems,
    such as compressions and repeats, but not necessarily within known
    annotated human repeat sequence elements (e.g. Alu). Where the
    sequence is ambiguous, there is an annotation using the 'unsure'
    feature key.
    The following abbreviations are used to associate primary accession
    numbers given in the feature table with their source databases:
    Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
    on the WORMPEP database can be found at
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
    was generated from part of bacterial clone contigs of human
    chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
    Group. Further information can be found at
    http://www.sanger.ac.uk/HGP/Chr1
    RP11-54H19 is from the library RPCI-11.1 constructed at the Roswell
    Park Cancer Institute by the group of Pieter de Jong. For further
    details see http://bacpac.med.buffalo.edu/
    VECTOR: pBACe3.6
    This sequence is the entire insert of clone RP11-17216 The true
    right end of clone RP11-17216 is at 50458 in this sequence.
FEATURES   Location/Qualifiers
 source   1..164168
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP11-54H19"
    /clone_lib="RPCI-11.1"

ORIGIN

  Query Match 85.2%; Score 23; DB 9; Length 164168;
  Best Local Similarity 100.0%; Pred. No. 16;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCGGGGTGAACGGGG 23
    |||||||||||||||||||||||
Db 129362 GGTGACTCACCGGGGTGAACGGGG 129384

RESULT 14
LOCUS   AC007227 164179 bp DNA linear PRI 20-NOV-1999
DEFINITION Homo sapiens chromosome 1 clone RPCI-11_54H19, complete sequence.
ACCESSION AC007227
VERSION AC007227.3 GI:6456148
KEYWORDS HTG.
SOURCE   Homo sapiens (human)
 ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164179)
 AUTHORS Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
    Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
    Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
    Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
    Han,C. and Deaven,L.
 TITLE   Sequencing of Human Chromosome 1
 JOURNAL Unpublished
REFERENCE 2 (bases 1 to 164179)
 AUTHORS Ricke,D.O.
 TITLE   Large Scale Sequence Analysis and Annotation with the Sequence
    Comparison Analysis (SCAN) System
 JOURNAL Unpublished
REFERENCE 3 (bases 1 to 164179)
 AUTHORS Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
    Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
    Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
    Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
    McMurry,K., Han,C. and Deaven,L.
 TITLE   Direct Submission
 JOURNAL Submitted (06-APR-1999) Center for Human Genome Studies, DOE Joint
    Genome Institute, Los Alamos National Laboratory, MS M888, Los
    Alamos, NM 87545, USA
REFERENCE 4 (bases 1 to 164179)
 AUTHORS Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
    Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
    Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
    Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
    McMurry,K., Han,C. and Deaven,L.
 TITLE   Direct Submission
 JOURNAL Submitted (20-NOV-1999) Center for Human Genome Studies, DOE Joint
    Genome Institute, Los Alamos National Laboratory, MS M888, Los
    Alamos, NM 87545, USA
COMMENT   On Nov 20, 1999 this sequence version replaced gi:5103801.
FEATURES   Location/Qualifiers
 source   1..164179
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RPCI-11_54H19"
 misc_feature join(1803..1889,3009..3122,4079..4159,8103..8192,
    8296..8385,8783..9025,9799..9939,10100..10699,
    12907..13098,15854..16030,17315..17536,23403..23561,
    23978..24068,25002..25079,25486..25623,25808..25891)
    /note="KIAA protein (Homo sapiens) BAA83041.1"
 repeat_region 2627..2766
    /rpt_family="Alu"
 repeat_region 44..354
    /rpt_family="Alu"
 repeat_region 175..474
    /rpt_family="Alu"
 repeat_region 366..665
    /rpt_family="Alu"
 repeat_region 921..1071
    /rpt_family="MIR"
 repeat_region ...
    /rpt_family="Alu"

```
repeat_region   3617..3907
                /rpt_family="Alu"
repeat_region   4568..4914
                /rpt_family="Alu"
repeat_region   complement(5455..5755)
                /rpt_family="Alu"
repeat_region   complement(6422..7051)
                /rpt_family="Alu"
repeat_region   7069..7507
                /rpt_family="Alu"
repeat_region   7541..7717
                /rpt_family="L1"
repeat_region   9550..9679
                /rpt_family="MIR"
repeat_region   complement(11237..11737)
                /rpt_family="Alu"
repeat_region   11961..12099
                /rpt_family="MIR2"
repeat_region   complement(12187..12512)
                /rpt_family="Alu"
repeat_region   complement(13983..14286)
                /rpt_family="Alu"
repeat_region   complement(14310..14594)
                /rpt_family="Alu"
repeat_region   2077..20164
                /rpt_family="Alu"
repeat_region   complement(20560..20876)
                /rpt_family="Alu"
repeat_region   complement(20930..21341)
                /rpt_family="Alu"
repeat_region   complement(21364..21673)
                /rpt_family="Alu"
misc_feature    15109..15184
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   complement(14590..14933)
                /rpt_family="Alu"
repeat_region   21737..22047
                /rpt_family="Alu"
repeat_region   16650..16947
                /rpt_family="Alu"
repeat_region   complement(17672..17965)
                /rpt_family="Alu"
repeat_region   19255..19569
                /rpt_family="Alu"
misc_feature    complement(join(22972..23038,23401..23562,23974..
                24053))
                /note="NCI_CGAP Kidil Homo sapiens cDNA (AW000913)"
misc_feature    join(23974..24069,25001..25079,25481..25537)
                /note="EST28865 Cerebellum II Homosapiens cDNA (AA325805)"
misc_feature    complement(25847..25930)
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    27587..27873
                /note="GRAIL 2 excellent exon, frame 1"
repeat_region   30267..30394
                /rpt_family="Alu"
misc_feature    join(30610..30686,31242..31321,31559..31721,32071..32187)
                /note="Stratagene schizo brain sll Homo sapiens
                (AA663560)"
misc_feature    complement(33070..34744)
                /note="Human gene for bone gla protein (BGP). Accession
                number X04143"
misc_feature    complement(join(33322..33453,33649..33720))
                /note="Human osteocalcin precursor (gamma-carboxyglutamic
                acid). Accession number P02818"
misc_feature    complement(34258..35135)
                /note="Human osteocalcin gene, 5' end, and promoter region
                (M34013)"
repeat_region   complement(35719..36177)
                /rpt_family="Alu"
misc_feature    complement(join(36901..37900,39601..40600,41345..41500,
                42267..42400,45881..46000,49501..49782,52201..53200))
                /note="Homo sapiens polyamine modulated factor-1
                (AF141309)"
repeat_region   38224..38381

repeat_region   38448..38781
                /rpt_family="L1"
repeat_region   38735..39038
                /rpt_family="Alu"
repeat_region   39252..39367
                /rpt_family="Alu"
repeat_region   39395..39520
                /rpt_family="Alu"
repeat_region   40725..41022
                /rpt_family="L1"
repeat_region   41058..41317
                /rpt_family="Alu"
repeat_region   41593..41667
                /rpt_family="Alu"
repeat_region   41673..41977
                /rpt_family="MIR2"
misc_feature    complement(42789..42889)
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   complement(42395..42527)
                /rpt_family="Alu"
repeat_region   42562..42704
                /rpt_family="Alu"
misc_feature    complement(43007..43300)
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   43395..43782
                /rpt_family="Alu"
misc_feature    complement(44092..44197)
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    complement(44770..45050)
                /note="GRAIL 2 excellent exon, frame 1"
repeat_region   complement(45070..45248)
                /rpt_family="Alu"
repeat_region   complement(45264..45564)
                /rpt_family="Alu"
repeat_region   complement(45567..45701)
                /rpt_family="Alu"
repeat_region   46611..46910
                /rpt_family="Alu"
repeat_region   47258..47576
                /rpt_family="Alu"
misc_feature    complement(48074..48187)
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   48208..48545
                /rpt_family="Alu"
repeat_region   48512..48678
                /rpt_family="Alu"
repeat_region   48669..48996
                /rpt_family="Alu"
repeat_region   complement(49023..49273)
                /rpt_family="Alu"
repeat_region   49717..49867
                /rpt_family="Alu"
repeat_region   50215..50644
                /rpt_family="MIR"
repeat_region   51016..51174
                /rpt_family="L1"
repeat_region   51301..51581
                /rpt_family="Alu"
repeat_region   complement(51693..51987)
                /rpt_family="Alu"
repeat_region   complement(52848..53012)
```

```
Query Match            85.2%;  Score  23;  DB  9;  Length 164179;
Best Local Similarity 100.0%;  Pred. No. 16;
Matches   23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GGTGACTCACCGGCGGTGAACCGGGG 23
            ||||||||||||| ||||||||||||
Db    34807 GGTGACTCACCGGGGGTGAACCGGGG 34785

RESULT 15
```

```
LOCUS       BX511245               95687 bp    DNA     linear   VRT 21-JUL-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-125O16 in linkage group 17,
            complete sequence.
ACCESSION   BX511245
VERSION     BX511245.5  GI:50470853
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 95687)
  AUTHORS   Phillimore,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 21, 2004 this sequence version replaced gi:50080055.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU). For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-125O16 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1..95687
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-125O16"
                     /clone_lib="CHORI-211"
ORIGIN
```

```
Query Match      74.8%;  Score 20.2;  DB 5;  Length 95687;
Best Local Similarity  88.0%;  Pred. No. 2.7e+02;
Matches 22;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;

Qy    3 TGACTCACCGGGGTGAACGGGGCATT 27
          |||||||| ||||| |||| ||||||
Db 10650 TGACACACCGTGTGAATGGGGCATT 10674
```